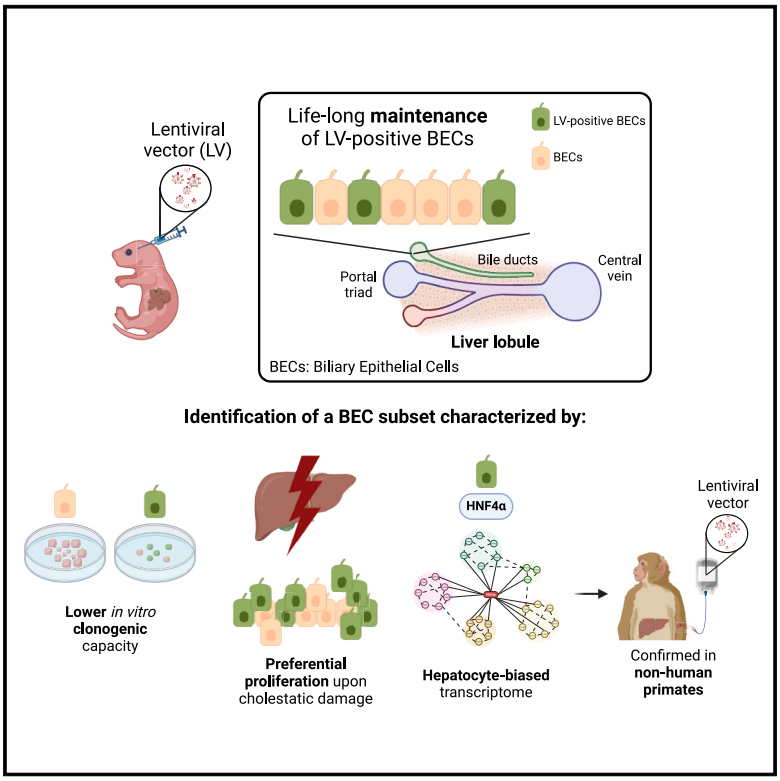
Cell Reports

Article

Identification of hepatocyte-primed cholangiocytes in the homeostatic liver by *in vivo* lentiviral gene transfer to mice and non-human primates

Graphical abstract



Authors

Michela Milani, Francesco Starinieri, Anna Fabiano, ..., Meritxell Huch, Luigi Naldini, Alessio Cantore

Correspondence

cantore.alessio@hsr.it

In brief

Milani et al. identify hepatocyte-primed liver biliary epithelial cells (BECs) in homeostatic conditions that are endowed with regenerative potential. They captured this BEC heterogeneity according to permissiveness to *in vivo* lentiviral gene delivery and the hepatocyte-biased transcriptome in mice and non-human primates (NHPs), highlighting the potential for lifelong and disease-resistant corrective gene therapy.

Highlights

- Heterogeneity of liver BECs in mice and NHPs in the homeostatic liver
- A BEC subset characterized by hepatocyte-biased transcriptional profile orchestrated by HNF4α
- Preferential proliferation of hepatocyte-primed BECs upon cholestatic liver damage
- Long-term gene transfer into BECs by *in vivo* lentiviral vector administration

Milani et al., 2025, Cell Reports *44*, 115341
March 25, 2025
https://doi.org/10.1016/j.celrep.2025.115341

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Identification of hepatocyte-primed cholangiocytes in the homeostatic liver by *in vivo* lentiviral gene transfer to mice and non-human primates

Michela Milani,[1] Francesco Starinieri,[1,9] Anna Fabiano,[1,9] Stefano Beretta,[1] Tiziana Plati,[1] Cesare Canepari,[1,2] Mauro Biffi,[1] Fabio Russo,[1] Valeria Berno,[4] Rossana Norata,[1] Francesca Sanvito,[1,3] Ivan Merelli,[5] Luigi Aloia,[6,8] Meritxell Huch,[6,7] Luigi Naldini,[1,2] and Alessio Cantore[1,2,10,*]

[1]San Raffaele Telethon Institute for Gene Therapy, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy
[2]Vita Salute San Raffaele University, 20132 Milan, Italy
[3]Pathology Unit, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy
[4]Advanced Light and Electron Microscopy BioImaging Center (ALEMBIC), IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy
[5]Institute for Biomedical Technologies, National Research Council, 20054 Segrate (MI), Italy
[6]The Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK
[7]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[8]Present address: Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge CB2 0AA, UK
[9]These authors contributed equally
[10]Lead contact
*Correspondence: cantore.alessio@hsr.it
https://doi.org/10.1016/j.celrep.2025.115341

SUMMARY

Liver regeneration is supported by hepatocytes and, in certain conditions, biliary epithelial cells (BECs). BECs are facultative liver stem cells that form organoids in culture and engraft in damaged livers. However, BEC heterogeneity in the homeostatic liver remains to be fully elucidated. Here, we exploit systemic lentiviral vector (LV) administration to achieve efficient and lifelong gene transfer to BECs in mice. We find that LV-marked BECs retain organoid formation potential and predominantly respond to liver damage; however, they are less clonogenic and display a hepatocyte-primed transcriptome compared to untransduced BECs. We thus identify a BEC subset committed to hepatocyte lineage in the absence of liver damage, characterized by a transcriptional network orchestrated by hepatocyte nuclear factor 4α. We also report *in vivo* targeting of such BECs in non-human primates. This work highlights intrinsic BEC heterogeneity and that *in vivo* LV gene transfer to the liver may persist following BEC-mediated repair of hepatic damage.

INTRODUCTION

The liver has an extensive regenerative capacity, as testified by its ability to regenerate following partial hepatectomy and reconstruct an entire liver after transplantation of a small portion.[1] Despite being fully differentiated cells, hepatocytes maintain self-renewal capacity and the potential to engraft and repopulate a damaged host liver and, thus, can behave as stem cells in case of need.[2,3] Other liver cell types have been shown to possess stem-cell-like features, as they contribute to liver regeneration according to the type and extent of tissue damage.[4] Such cells are referred to as facultative stem cells since they can complement hepatocytes in liver regeneration, particularly when the capacity of the latter is hampered by the extent and/or duration of damage.[5] Cholangiocytes, also referred to as biliary epithelial cells (BECs), have been reported to be self-renewing, facultative stem cells able to generate *ex vivo* liver organoids that can be expanded in culture and engraft in a damaged liver.[6,7] These cells have also been shown to respond to damage by proliferation in a process known as ductular reaction and transdifferentiate into hepatocytes when the proliferation of the latter ones is impaired.[8–11] In addition, cholangiocytes' plasticity has been highlighted by transplanting organoids derived from human intra- or extra-hepatic bile ducts, showing that these cells can acquire the phenotype of intrahepatic biliary epithelium independently of their initial origin.[12] Moreover, Aizarani et al. described heterogeneity among human intrahepatic cholangiocytes by showing that only a fraction of these cells give rise to organoids and that there are subsets with different organoid-forming potential.[13] However, intrinsic BEC heterogeneity in the liver at a steady state and its role in regeneration remains to be fully elucidated. The regenerative capacity of the liver is an important feature in the context of gene therapy since efficacy may be maintained even after tissue damage, provided that the cells supporting this process are efficiently targeted. Thus, gaining a better understanding of BEC biology also has important therapeutic implications. Genetic modification of hepatocytes can indeed be exploited for therapeutic purposes in several inherited and acquired diseases, such as inherited metabolic disorders and clotting factor deficiency hemophilia.[14] Successful results

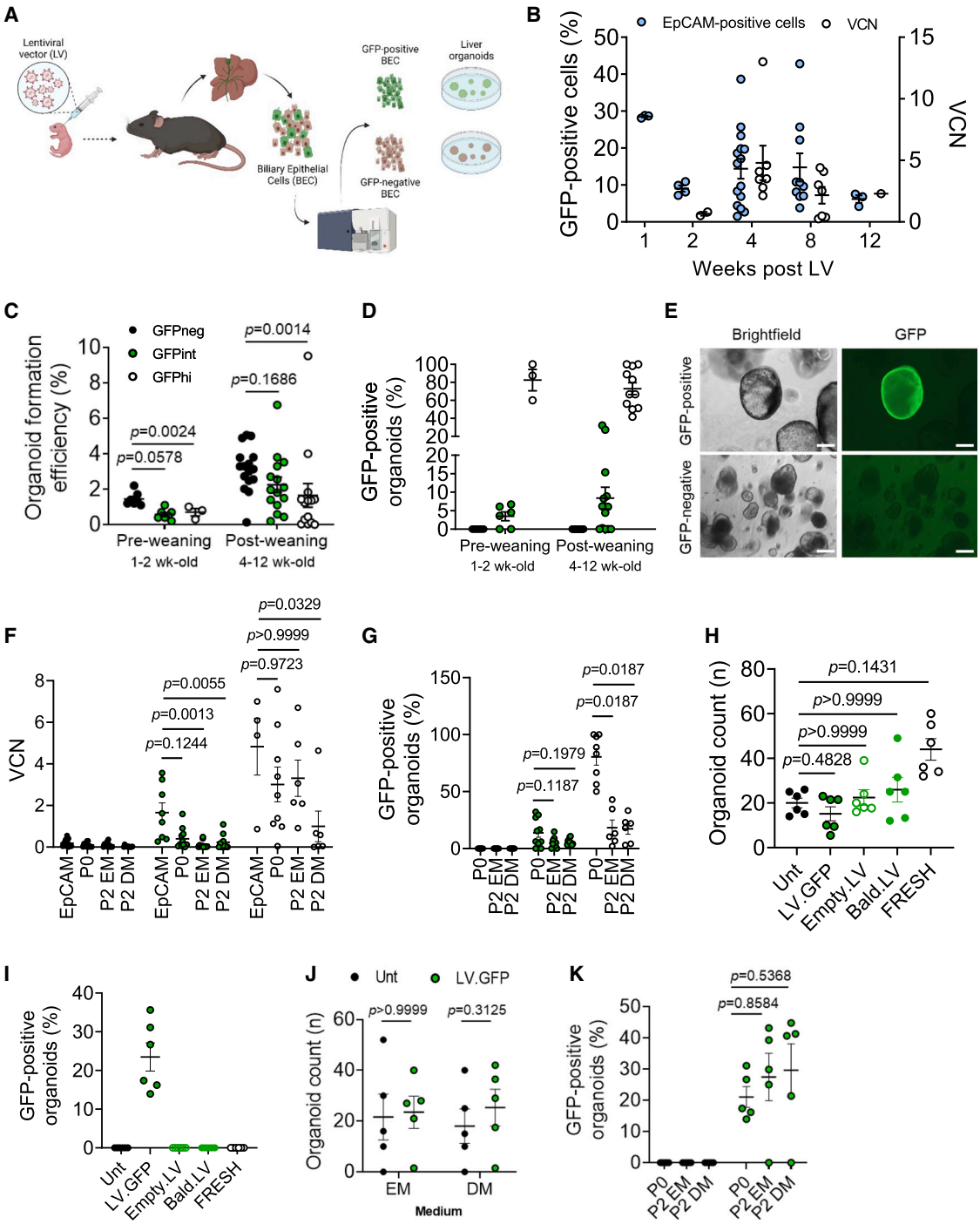


Figure 1. Efficient transduction of intrahepatic cholangiocytes by *in vivo* LV administration to newborn mice

(A) Experimental scheme: LV encoding for a ubiquitously expressed GFP is administered *in vivo* intravenously (i.v.) in the temporal vein of newborn mice at 5 × 10^{10} TU/Kg. The liver is collected at different time points after LV administration, and EpCAM-positive bile epithelial cells (BECs) are isolated and sorted according to GFP expression. Liver organoids are then derived from GFP-positive or GFP-negative (GFPneg) BECs.

(B) Single values and mean with standard error of the mean (SEM) of GFP-positive EpCAM-positive cells (left axis) or vector copies per diploid genome (vector copy number [VCN], right axis) measured in GFP-positive EpCAM-positive sorted cells at the indicated time points after LV *in vivo* administration from mice treated as newborns (5 × 10^{10} TU/Kg); 1 week $n = 3$, 2 weeks $n = 4$, 4 weeks $n = 15$, 8 weeks $n = 10$, and 12 weeks $n = 3$. Pool of 7 independent experiments.

(C and D) Single values and mean with SEM of the organoid formation efficiency (C) or GFP-positive organoids (D) of GFPneg, GFP-intermediate (GFPint), or GFP-high (GFPhi) EpCAM-positive sorted cells according to the age of mice at the moment of liver collection; pre-weaning: GFPneg $n = 7$, GFPint $n = 7$, and GFPhi $n = 3$; post-weaning: GFPneg $n = 16$, GFPint $n = 15$, and GFPhi $n = 14$. Kruskal-Wallis test with Dunn's multiple comparisons test.

(legend continued on next page)

have been obtained in hemophilia with liver gene therapies that have recently reached marketing authorization. These approaches are based on a single intravenous (i.v.) administration of adeno-associated viral (AAV) vectors delivering a functional copy of a coagulation factor gene in adult humans affected by hemophilia.[15,16] However, the non-integrating nature of AAV vectors poses challenges for the persistence of the therapeutic transgene following hepatocytes' proliferation, such as in postnatal liver growth and homeostatic turnover, but also in case of damage-driven proliferation. This aspect is particularly relevant in the context of metabolic diseases causing damage to the liver. For this reason, the use of lentiviral vectors (LVs) stably integrating their genetic cargo into target cell chromatin may be advantageous. We have previously reported efficient and stable gene transfer to the liver in mice, dogs, and non-human primates (NHPs) following i.v. administration of LVs.[17–19] Here, we investigated the efficiency and stability of LV-mediated *in vivo* gene transfer into cholangiocytes in mice and NHPs. Moreover, we evaluated the organoid formation efficiency, transcriptomic profile, response to damage, and engrafting potential of LV-transduced and untransduced BECs. We provide evidence of intrinsic heterogeneity of cholangiocytes in the homeostatic liver, highlighting a more hepatocyte-primed and less clonogenic population that is highly permissive to gene transfer and characterized by a transcriptional network orchestrated by hepatocyte nuclear factor 4α (HNF4α).

RESULTS

Efficient and stable gene transfer to cholangiocytes by *in vivo* LV administration

We administered LVs that express GFP under the control of a ubiquitous phosphoglycerate kinase (PGK) promoter to newborn mice by i.v. delivery to broadly capture transduced cell types by GFP expression. i.v.-administered LVs efficiently transduce hepatocytes, liver sinusoidal endothelial cells, and Kupffer cells, as previously shown,[18,20] but their ability to reach and transduce cholangiocytes is unknown. We sorted transduced cells by fluorescence-activated cell sorting (FACS) at different times after LV administration and analyzed GFP expression in epithelial cell adhesion molecule (EpCAM)-positive BECs and measured LV copies per diploid genome (vector copy number [VCN]; Figure 1A). We observed an average of 10%–15% GFP-positive EpCAM-positive cells with a VCN between 1 and 5 (Figure 1B), indicating efficient *in vivo* gene transfer into intrahepatic BECs by LVs. We then assessed the organoid formation efficiency of GFP-positive or -negative ductal cells sorted according to GFP fluorescence intensity (GFP intermediate, GFP high). We observed that GFP-positive cells displayed lower organoid formation efficiency than the negative counterparts at all tested time points (Figure 1C). We analyzed GFP expression in organoids by fluorescence microscopy and observed that not all the organoids derived from the GFP-positive EpCAM-positive cells were GFP positive, suggesting preferential expansion of the few GFP-negative contaminating cells (Figures 1D and 1E). Moreover, when cultured in either expansion medium (EM) or differentiation medium (DM), organoids derived from LV-transduced cells were counterselected over time, as shown by a progressive reduction in both the VCN and percentage of GFP-positive organoids (Figures 1F and 1G). These data suggest that LV-transduced cholangiocytes display a lower clonogenic potential compared to their untransduced counterparts. We also confirmed this outcome when newborn mice were transduced with LVs in which GFP expression was driven by a previously described hepatocyte-specific enhancer/promoter (enhanced transthyretin [ET]). Interestingly, we observed that all transduced EpCAM-positive cells were GFP positive, as the sorted GFP-negative BECs did not show LV copies, thus indicating that this promoter is active in this cell type (Figures S1A–S1E). Similar to the phenotype observed above, whereas GFP-positive ductal cells showed similar organoid formation efficiency to GFP-negative counterparts, they gave rise to organoids only partly GFP positive and that were progressively counterselected in culture (see Figures S1B–S1E). We thus set out to investigate if this phenotype was a consequence of LV transduction or due to a cell-intrinsic feature of a cholangiocyte subset. To this end, we transduced EpCAM-positive cells from adult untreated mice with LV.GFP *in vitro*. We included unmanipulated cells (FRESH), mock-transduced cells (centrifuged without LV, Unt), or cells exposed to empty.LV (without genome) or bald.LV (without envelope protein) as controls. We did not find differences in the organoid counts between the tested conditions except for a tendency for higher organoid counts in the FRESH condition (Figure 1H). As expected, the only GFP signal came from the LV.GFP samples (Figure 1I). Importantly, there was no difference in organoid counts or decrease in GFP-positive organoids during culture in either EM or DM (Figures 1J and 1K), contrary to what was observed when organoids were derived from BECs of LV-transduced mice (see Figures 1F and 1G). These data suggest that LV transduction and expression *per se* do not impact the organoid formation efficiency or self-renewal capacity of liver cholangiocytes. In line with this finding, we observed that the percentage of GFP-positive ductal cells and LV VCNs 1 year after LV administration remained in the same range as those observed at previous times of analysis (Figures S1F, S1G,

(E) Representative images of organoids derived from GFP-positive or GFPneg EpCAM-positive cells. Scale bar: 150μm.
(F) Single values and mean with SEM of VCN in GFPneg ($n = 11$), GFPint ($n = 10$), or GFPhi ($n = 8$) EpCAM-positive sorted cells, organoids (P0), or organoids kept in culture in expansion medium (EM) or differentiation medium (DM) after two passages (P2). Kruskal-Wallis test with Dunn's multiple comparisons test.
(G) Single values and mean with SEM of the percentage of GFP-positive organoids shown in (F). Kruskal-Wallis test with Dunn's multiple comparisons test.
(H and I) Single values and mean with SEM of number of organoids (H) or the percentage of GFP-positive organoids (I) derived from fresh EpCAM-positive cells (FRESH) or after spinoculation in the presence or absence (Unt) of LV (LV.GFP) or control LV produced without genome (empty.LV) or without envelope protein (bald.LV). Pool of 2 independent experiments; $n = 6$ mice. Friedman test with Dunn's multiple comparisons test (H).
(J and K) Single values and mean with SEM of organoid count (J) or the percentage of GFP-positive organoids (K) of organoids shown in (H) ($n = 5$) cultured in EM or DM analyzed at the indicated time points. Wilcoxon matched-pairs signed rank test (J). Friedman test with Dunn's multiple comparisons test (K).

and 1B), supporting nearly lifelong maintenance of LV-transduced BECs. The organoid formation efficiencies were similar between GFP-positive and -negative ductal cells at this time point (Figures S1H and S1I).

LV transduction highlights cholangiocyte heterogeneity

We hypothesized that LVs preferably transduce a more hepatocyte-committed cholangiocyte subset. This hypothesis may be supported by the observation that all LV-transduced EpCAM-positive cells expressed the hepatocyte-specific promoter (see Figure S1A). We investigated TROP2 as a candidate to enrich for a more hepatocyte-biased and less organoid-forming cholangiocyte subset based on a previous report related to human BECs.[13] We FACS-sorted EpCAM-positive cells from LV-transduced mice according to TROP2 expression or GFP expression into negative, intermediate, and high (Figures S2A and S2B). We observed that LV-transduced GFP-positive BECs were enriched in the TROP2-negative population (Figure S2C) and that the organoid formation efficiency was independent of TROP2 expression, thus suggesting that TROP2 expression does not discriminate between different cholangiocyte subsets in mice (Figures S2D and S2E). We performed a histological analysis of the livers of transduced mice and observed that not all the cells of each duct were GFP positive and that most GFP-positive ductal cells belonged to ducts larger than 15 μm.[21] These data suggest that LV transduction of ductal cells is unevenly distributed (Figures S2F and S2G). We also performed a histological analysis of cultured organoids. We observed dilated glandular-like structures lined by epithelial cells, with slight cellular heterogeneity, without substantial differences between organoids cultured in different media or derived from GFP-positive vs. GFP-negative BEC (Figure S2H). We then performed gene expression analysis on selected hepatocyte-specific genes (*Alb*, *HNF4a*, and *Cyp3a11*) or cholangiocyte-specific genes (*HNF1b*, *EpCAM*, *Krt7*, and *Krt17*) of sorted EpCAM-positive BECs and organoids derived from GFP-negative, GFP-intermediate, and GFP-high BECs, respectively (Figures S3A–S3G). Remarkably, we observed that *Alb* was highly expressed by FACS-sorted BECs, being higher in GFP-positive compared to -negative BECs. Similarly, *HNF1b* was more highly expressed by GFP-positive than -negative BECs. *HNF1b* was previously described to have heterogeneous expression among BECs located in the portal triad compared to the canals of Hering.[22] Moreover, we observed that hepatocyte-specific genes tended to be upregulated following culture in DM, while the opposite was true for cholangiocyte-specific genes. Interestingly, organoids deriving from GFP-high compared to -negative cholangiocytes showed significantly higher expression of *Alb* and *HNF4a* before the first passage and *Alb* and *Cyp3a11* following culture in DM (Figures S3H–S3M). To assess the heterogeneity of *Alb* and *Hnf4a* expression, we manually picked single GFP-positive or GFP-negative organoids and analyzed them separately, thus also avoiding the contaminating GFP-negative organoids present in those derived from GFP-positive BECs (see Figure 1E). We observed that, on average, the GFP-positive BEC-derived organoids expressed significantly more *Alb* and *HNF4a* than those derived from the GFP-negative BECs (Figure 2A). To confirm gene expression data at the protein level, we analyzed albumin production at different time points in the supernatant of cultured organoids grown from GFP-positive or -negative BECs, and we performed co-staining of GFP and HNF4α or HNF1β by consequential immunofluorescence and immunohistochemistry analysis of the same liver slice. Interestingly, we found higher albumin production by organoids grown from GFP-positive BECs after culturing them in EM or DM (Figure 2B). In addition, despite few BECs being HNF4α positive, we found a high frequency of HNF4α-positive BECs among GFP-positive BECs, and in accordance with gene expression data, most GFP-positive BECs were also HNF1β positive (Figures 2C and 2D). These data prompted us to perform a deep transcriptomic analysis of EpCAM-positive GFP-positive or EpCAM-positive GFP-negative cells 8 weeks following neonatal LV delivery. We found a remarkably different gene expression profile between the 2 populations with 874 differentially expressed genes (false discovery rate [FDR] < 0.01, absolute log fold change [abs(logFC)] > 2; Figures 3A and 3B). Post hoc analyses of genes upregulated in GFP-positive cells revealed a prevalent hepatocyte signature and highlighted biological processes typical of hepatocytes, such as cholesterol metabolism, coagulation cascade, bile secretion, and amino acid metabolism, indicating the intrinsic priming of these cholangiocytes toward a hepatocyte lineage (Figure 3C). Conversely, enriched categories of genes downregulated in GFP-positive cells showed other epithelial cell types, including cholangiocytes, among the first hits (Figure S4). We re-analyzed a previously published single-cell RNA sequencing (scRNA-seq) dataset of human livers to map the human orthologs of the genes that we found upregulated in GFP-positive BECs.[13] Interestingly, this gene signature was upregulated in those BECs expressing asialoglycoprotein receptor 1 (*ASGR1*), which were reported to have a fate bias toward the hepatocyte lineage[13] (Figures 3D, 3E, and S5A). These data prompted us to perform scRNA-seq analysis on two NHP liver samples. Importantly, we were able to find the signature identifying hepatocyte-primed BECs in a subset of cholangiocytes in both LV-treated and untreated NHPs, supporting the existence of this population at a steady state in the liver independent of LV administration (Figures 3F and 3G). These cells were also characterized by high expression of albumin, in line with the mouse data (Figure S5B). Network analysis of the genes upregulated in GFP-positive murine ductal cells highlighted a high level of interconnections (Figure S5C). Further analysis of these genes according to databases of transcription factors identified HNF4α as a master regulator of this transcriptional program based on a specific signature among all differentially expressed genes (Figure S5D).

Albumin-expressing cholangiocytes show a hepatocyte-primed transcriptome

To confirm the presence of a hepatocyte-primed cholangiocyte subset in the steady-state liver independent of LV transduction, we exploited a mouse line expressing an inducible Cre recombinase under the control of the *Alb* promoter (Alb-CreERT2), which has been shown to be expressed by BECs (see Figure S3A). We crossed these mice with a reporter mouse line in which one out of four different fluorescent transgenes is expressed upon Cre recombination[23] (Figure 4A). We found that, on average, 24% of EpCAM-positive cells were also positive for one Confetti

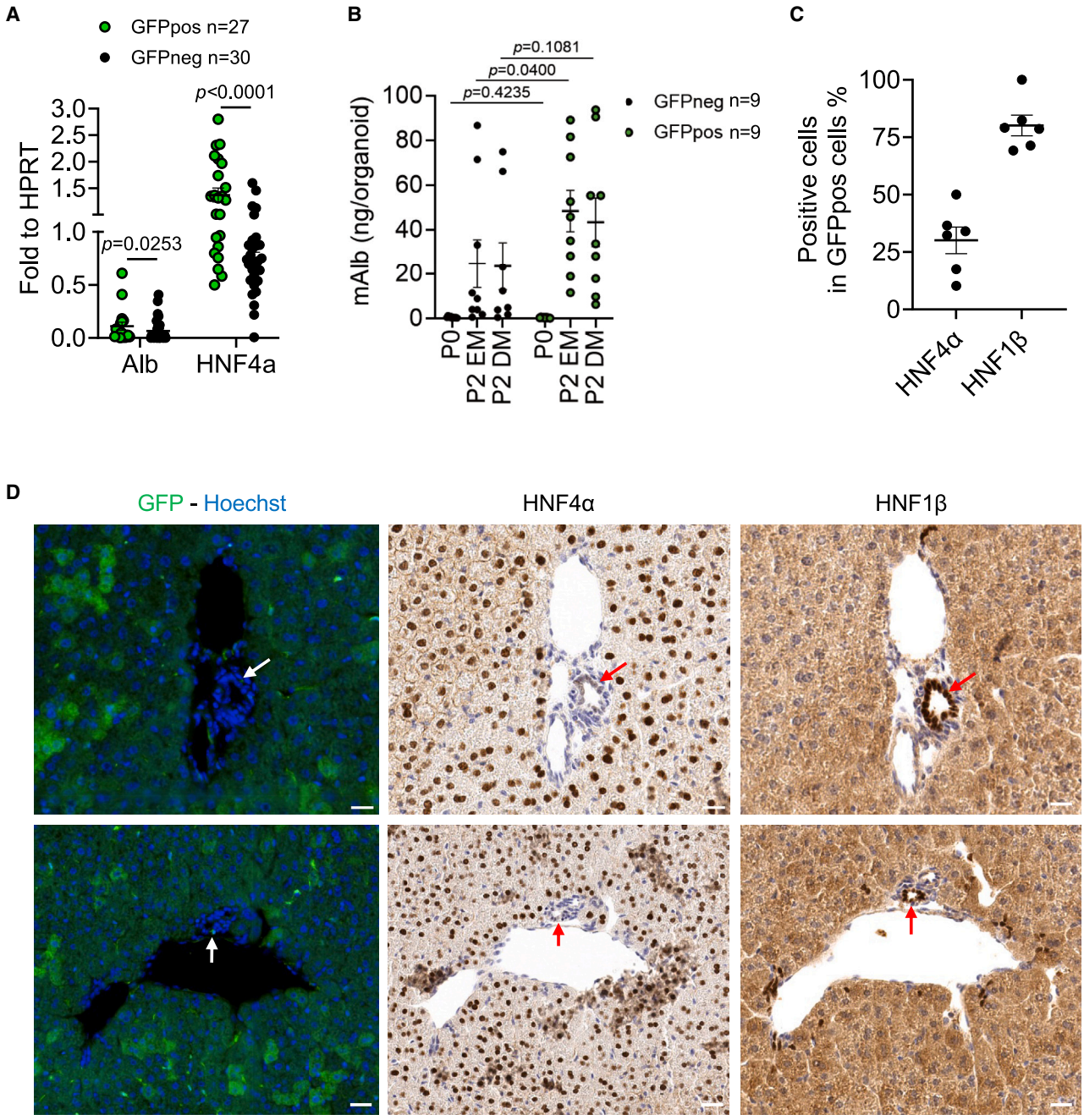


Figure 2. GFP-positive BECs generate organoids producing high amounts of albumin and are enriched for HNF4α and HNF1β markers

(A) Single values and mean with SEM of the expression of the indicated genes in picked GFP-positive (GFPpos; *n* = 27) or GFPneg (*n* = 30) single organoids obtained from EpCAM-positive BECs isolated from 8-week-old mice (*n* = 3) treated as newborns with i.v. administration of LVs expressing GFP under the control of the ubiquitous PGK promoter. Mann-Whitney test.

(B) Single values and mean with SEM of ng of murine albumin (mAlb) produced by organoids derived from GFPneg (*n* = 9) or GFPpos (*n* = 9) organoids (P0) or organoids kept in culture in expansion medium (EM) or differentiation medium (DM) after two passages (P2). Mann-Whitney test.

(C) Single values and mean with SEM of HNF4α- or HNF1β-positive BECs measured by immunohistochemistry, among GFPpos BECs identified by immunofluorescence on the same slice of liver, analyzed 8 weeks post-LV administration to newborn mice (*n* = 6).

(D) Representative images of GFPpos BECs positive for HNF4α or HNF1β marker as indicated (white or red arrows). Scale bar: 25 μm.

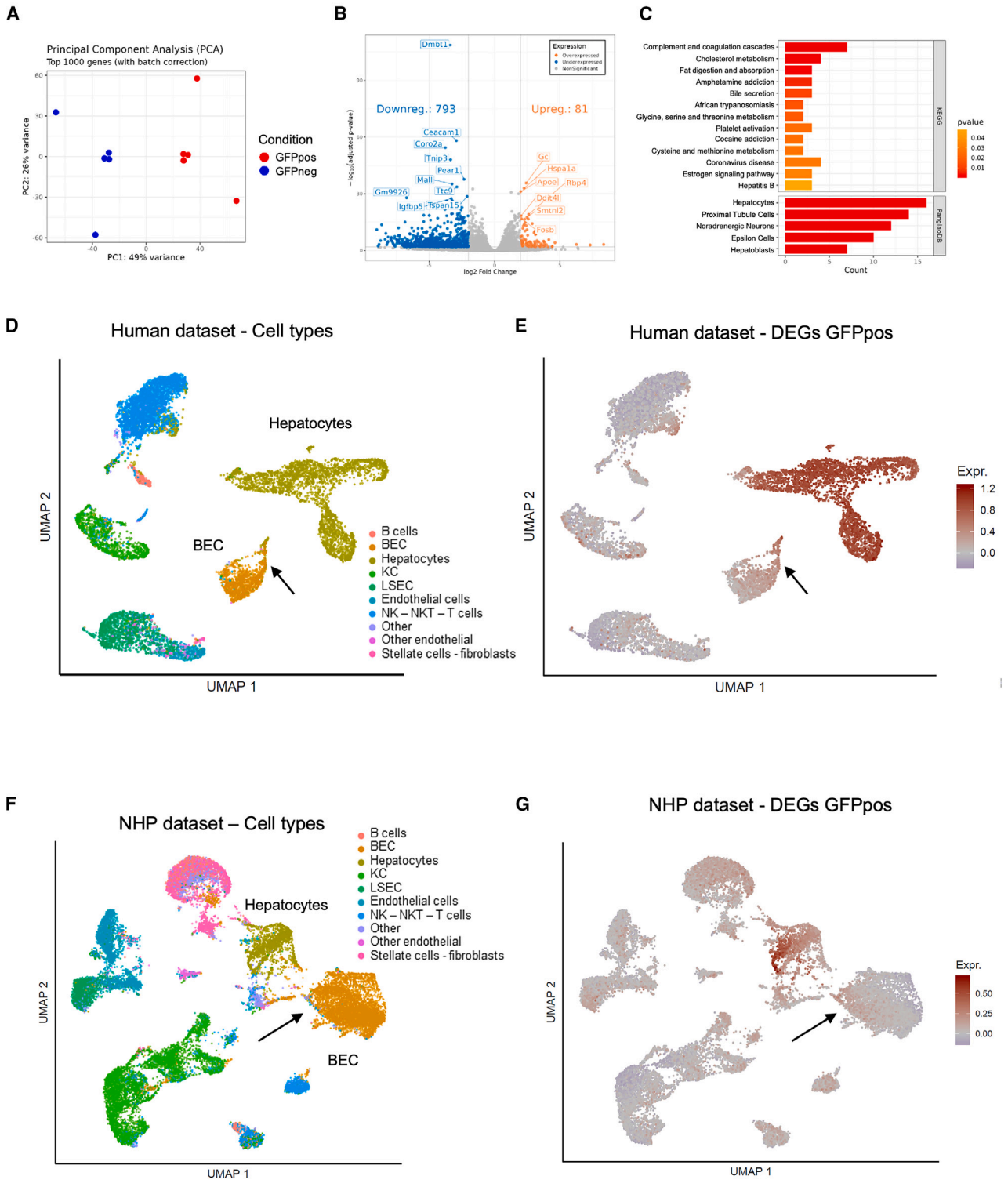


Figure 3. Transcriptomic analyses reveal a subpopulation of cholangiocytes primed toward hepatocyte lineage

(A) Principal-component analysis (PCA) on 1,000 genes obtained from transcriptomic analysis of GFP-positive (GFPpos; red) vs. GFP-negative (GFPneg; blue) EpCAM-positive cells isolated and sorted from 8-week-old mice treated by *in vivo* LV administration as newborns (n = 5, 5 × 10^{10} TU/Kg).

(B) Volcano plot showing the results of the differential expression analysis reporting log2 fold change on the x axis and $-\log_{10}$ of the adjusted p value on the y axis. Positive/negative differentially expressed genes (DEGs; FDR < 0.01, abs(logFC) > 2) are colored in orange/blue, respectively.

(legend continued on next page)

marker, indicating that these cells expressed the recombinase under the *Alb* promoter enough to recombine the Confetti construct (Figures 4B and S6A). Importantly, we did not detect any Confetti-positive cells in tamoxifen-untreated mice (Figure S6B). We then FACS-sorted EpCAM-positive, Confetti-positive, or Confetti-negative BECs. We observed lower organoid formation efficiency from the former than the latter, as well as a progressive decrease in Confetti-positive organoids in culture (Figures 4C and 4D). We performed gene expression analysis on hepatocyte-specific or cholangiocyte-specific genes, as above, and observed high *Alb* expression in sorted EpCAM cells, consistent with what was observed before (see Figure S3A), supporting the observed Alb-mediated recombination of the Confetti locus. In addition, we observed higher hepatocyte-specific gene expression in organoids derived from Confetti-positive compared to -negative ductal cells (Figures S6C–S6H). Since the phenotype of the organoids derived from Confetti-positive ductal cells was similar to that of those derived from the GFP-positive *in vivo* LV-transduced cells (see Figures 1C–1G and S3), we performed a transcriptomic analysis of the Confetti-positive and -negative ductal cells. Unsupervised clustering of these samples together with BECs isolated from LV.GFP-transduced mice (from Figure 3) showed similar transcriptional profiles of Confetti-positive and GFP-positive BECs, while Confetti-negative clustered with GFP-negative BECs (Figure 4E). Taken together, these data show that cholangiocytes with a more hepatocyte-primed transcriptome can be identified by both higher permissiveness to *in vivo* LV transduction and high *Alb* expression in an LV-independent manner and that they display a similar transcriptomic profile.

LV-transduced cholangiocytes predominantly expand upon liver cholestatic damage

To verify the capacity of LV-transduced BECs to respond to liver damage, we transduced mice as newborns with LV.GFP and then administered to some of them a toxic diet containing 0.1% of 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC), which is known to induce cholestatic damage.[8] One week later, we discontinued the diet and collected livers 2 weeks after for flow cytometry and molecular analyses (Figure 5A). We confirmed the liver damage in DDC-fed mice, as shown by weight loss and an increase in circulating liver enzymes, which gradually returned to the normal range (Figures 5B–5E). LV VCNs in the total liver were similar among DDC-fed and control mice, confirming similar transduction of the liver by LVs (Figure 5F). As expected, the percentage of EpCAM-positive cells increased, indicating ductular reaction in response to the liver damage (Figure 5G). Interestingly, GFP-positive BECs expanded 3-fold in DDC-fed compared to control mice, indicating a proliferative response of GFP-positive LV-transduced BECs (Figure 5H), while there was no difference in the percentage of GFP-positive cells in the remaining non-parenchymal cells (nPCs; Figure 5I). These data show that the subset of BECs identified and tracked here by LV transduction predominantly proliferates in response to liver damage and might better contribute to regeneration.

In vivo LV administration to NHPs transduces cholangiocytes that maintain their organoid-forming and engrafting potential

We then evaluated whether hepatic BECs can be transduced by systemic LV administration to NHPs, a more relevant model for biology in humans and from the perspective of clinical translation. We took advantage of livers collected from NHPs treated with LVs expressing human coagulation factor VIII (FVIII) or FIX transgenes from other studies.[19] Transgene expression was under the control of the hepatocyte-specific ET promoter. We processed the livers to isolate BECs according to a previously described protocol for human livers.[24] We obtained similar counts of organoids from both an untreated NHP and LV-treated NHPs at different LV doses (Figures 6A and 6B). We found LV VCNs in the isolated BECs and about 2% of single organoids with values directly proportional to the LV dose administered to the NHPs (Figures 6C and 6D). Note that the LV dose/kg administered to NHPs is 10- to 50-fold lower than that administered to mice since hepatocyte gene transfer by LV is more efficient in the former, as previously shown.[18,19] We detected human FVIII or FIX, according to the LV administered to the NHPs, in the supernatant of the organoids when cultured in either EM or DM (Figures 6E and 6F), indicating expression of the hepatocyte-specific, transthyretin-based ET promoter in both conditions, as shown above for murine organoids. We also detected similar amounts of NHP albumin produced by the organoids of LV-treated or untreated NHPs as a control for protein production and secretion (see Figure 6F). Note that, in this case, only 2% of the organoids derived from BECs isolated from LV-treated NHPs are actually LV transduced; thus, we do not expect to observe any difference in albumin production between these and organoids derived from BECs isolated from the untreated control NHP. We observed an increase in the expression of *Cyp3a4* in organoids cultured in DM compared to EM, while *Hnf4a* and *Krt7* did not significantly change (Figure 6G). Longitudinal gene expression analysis of cultured NHP organoids in EM showed a progressive increase in *Cyp3a4* expression, which was highest in DM, suggesting commitment toward the hepatocyte lineage even during culture in EM for a prolonged time (Figure 6H). FVIII was produced at similar levels at all times of analysis (Figure 6I). We assessed if the stiffness of the culture support impacted organoids' differentiation. We observed an

(C) Barplot of enriched categories (adjusted $p < 0.05$) on DEGs upregulated in GFPpos compared to GFPneg cells in the indicated databases. Counts on the x axis represent the number of enriched genes in the corresponding term, while bar colors represent the statistical significance (p value).
(D) Uniform manifold approximation and projection (UMAP) plot showing reanalysis of single-cell (sc)RNA-seq transcriptomes from healthy human livers from Aizarani et al.[13] Colors indicate the different cell types identified, as indicated. Arrow shows hepatocyte-biased BECs.
(E) UMAP plot showing expression of the DEGs upregulated in GFPpos EpCAM-positive cells identified in (B) mapped on the human dataset shown in (D).
(F) UMAP plot showing clustering of scRNA-seq transcriptomes from 2 NHP livers ($n = 1$ LV-treated NHP, $n = 1$ untreated NHP). Cells are colored according to the mapping of the cell types (clusters) from Aizarani et al.[13] Arrow shows hepatocyte-biased BECs.
(G) UMAP plot showing expression of the DEGs upregulated in GFPpos EpCAM-positive cells identified in (B) mapped on the NHP dataset shown in (F).

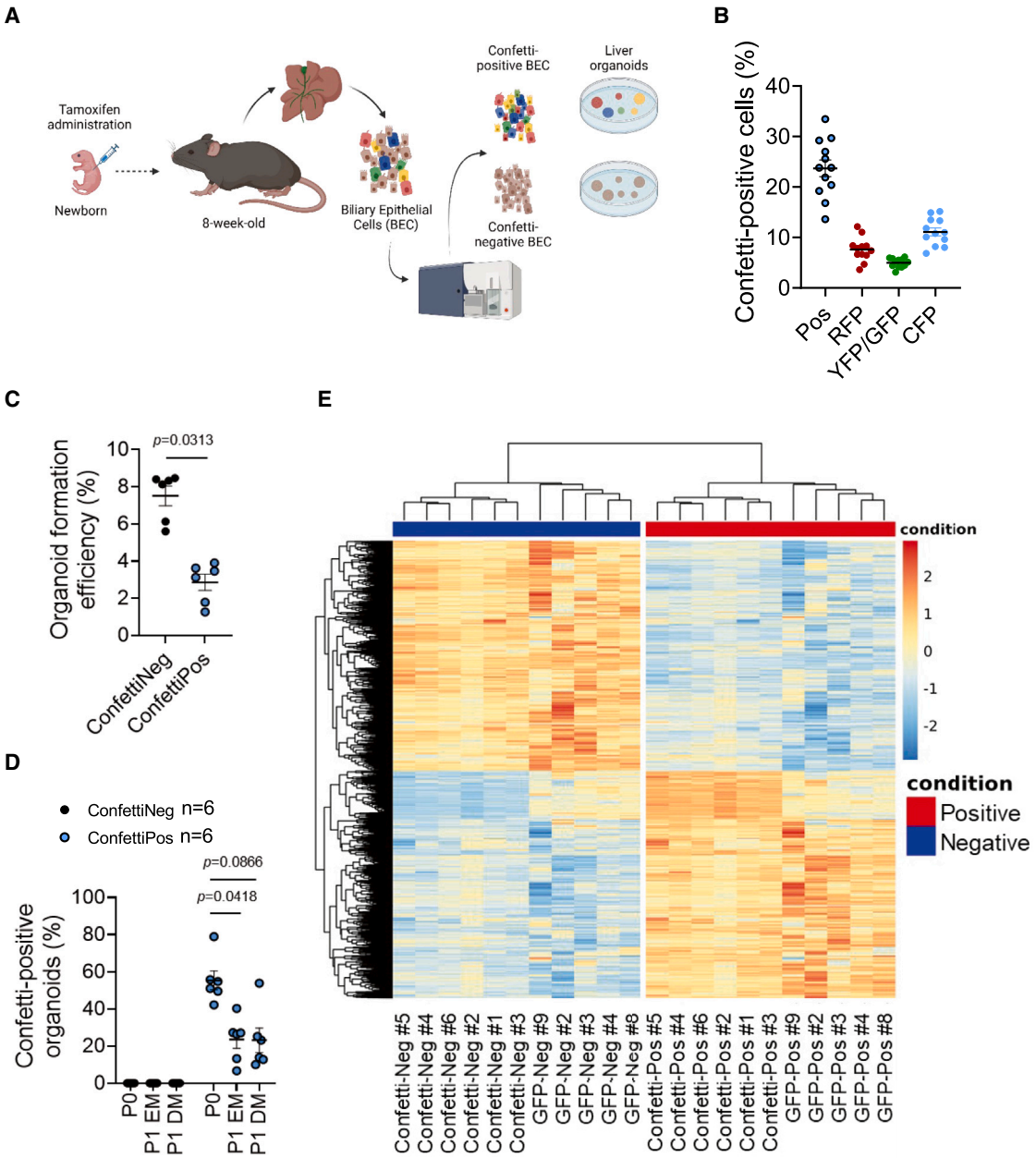


Figure 4. Marked cholangiocytes isolated from the liver of induced Alb-CreERT2/Confetti mice phenocopy GFP-positive cholangiocytes

(A) Experimental scheme: tamoxifen is administered *in vivo* subcutaneously to newborn mice. The liver is collected at 8 weeks of age, and EpCAM-positive BECs are isolated and sorted according to the expression of the Confetti markers. Liver organoids are then derived from Confetti-positive (ConfettiPos) or Confetti-negative (ConfettiNeg) BECs.

(B) Single values and mean with SEM of ConfettiPos EpCAM-positive cells isolated from the livers of 8-week-old Alb-CreERT2/Confetti mice ($n = 12$) administered with tamoxifen as newborns. Pos, Confetti-positive EpCAM-positive cells; RFP, red fluorescent protein; YFP/GFP, yellow/green fluorescent protein; CFP, cyan fluorescent protein.

(C and D) Single values and mean with SEM of the organoid formation efficiency (C) or the percentage of ConfettiPos organoids (D) of ConfettiNeg ($n = 6$) or ConfettiPos ($n = 6$) EpCAM-positive sorted cells at the indicated passages and cultured in EM or DM. Wilcoxon matched-pairs signed rank test (C). Friedman test with Dunn's multiple comparisons test (D).

(E) Heatmap of the expression of DEGs resulting from the comparison of positive (GFP or Confetti) samples with negative (GFP or Confetti) samples. Unsupervised clustering highlights the similarity of GFP (see Figure 3) and Confetti samples on both positive and negative conditions.

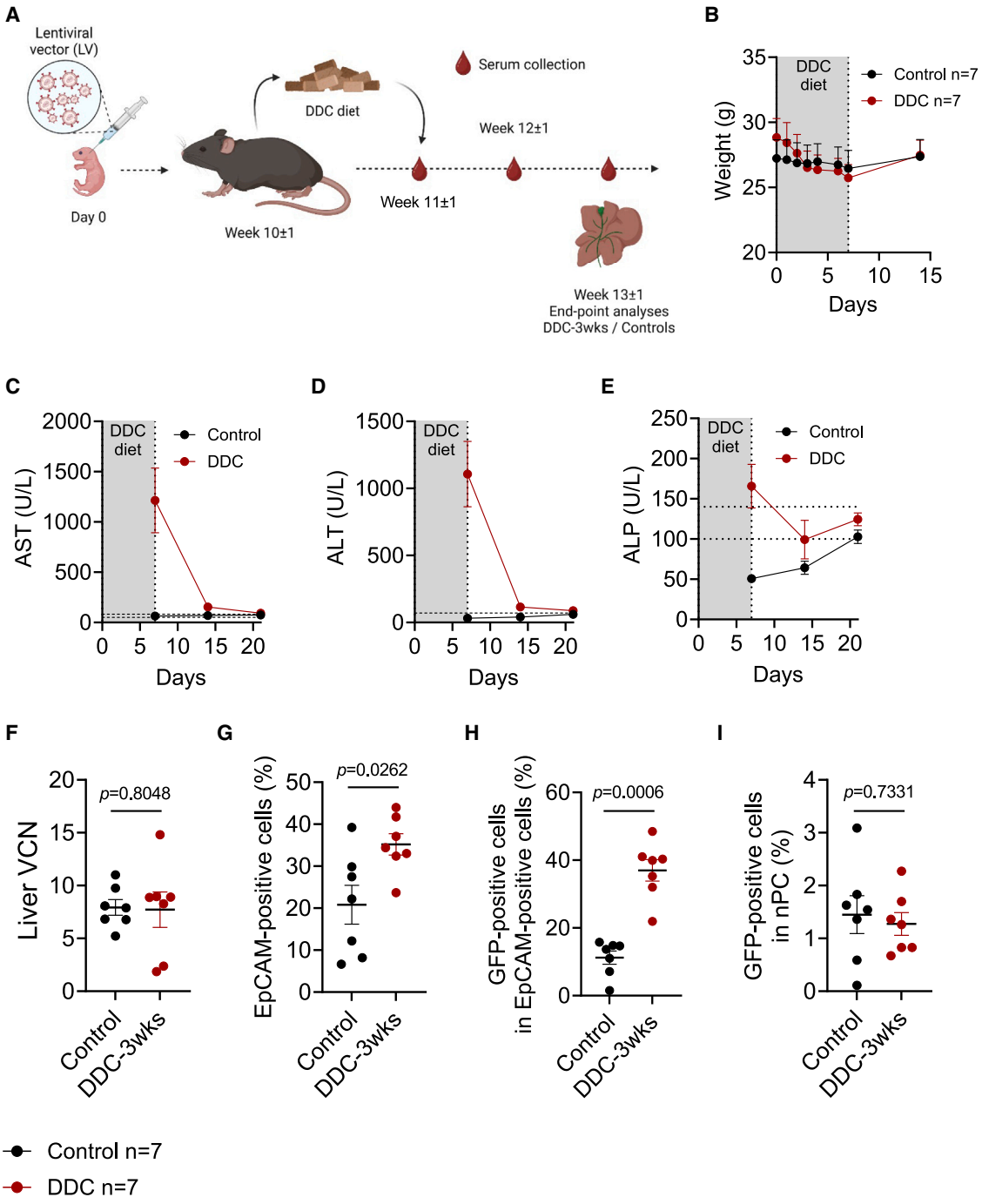


Figure 5. GFP-positive mouse cholangiocytes predominantly respond and expand upon cholestatic damage

(A) Experimental scheme: LV is administered i.v. to newborn mice (5×10^{10} TU/Kg) 9–10 weeks after cholestatic damage is induced by 0.1% 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) diet for 1 week, and proliferation of GFP-positive or GFP-negative BECs is analyzed at week 12–13 (2 weeks after the end of the DDC diet, $n = 7$ DDC-fed mice or $n = 7$ control mice).

(B–E) Mean with SEM of the weight (B) or serum concentration of aspartate aminotransferase (AST; C), alanine aminotransferase (ALT; D), or alkaline phosphatase (ALP; E) in DDC-fed or control mice measured at the indicated time points. The gray area represents the DDC diet time frame. Dashed lines: normal range.

(F) Single values and mean with SEM of VCN measured in the total liver of control or DDC-fed mice at the indicated time points. Mann-Whitney test.

(G–I) Single values and mean with SEM of the percentage of EpCAM-positive BECs (G), GFP-positive EpCAM-positive BECs (H), or GFP-positive non-parenchymal cells (nPCs; I) measured in the liver of control or DDC-fed mice at the indicated time point. Mann-Whitney test.

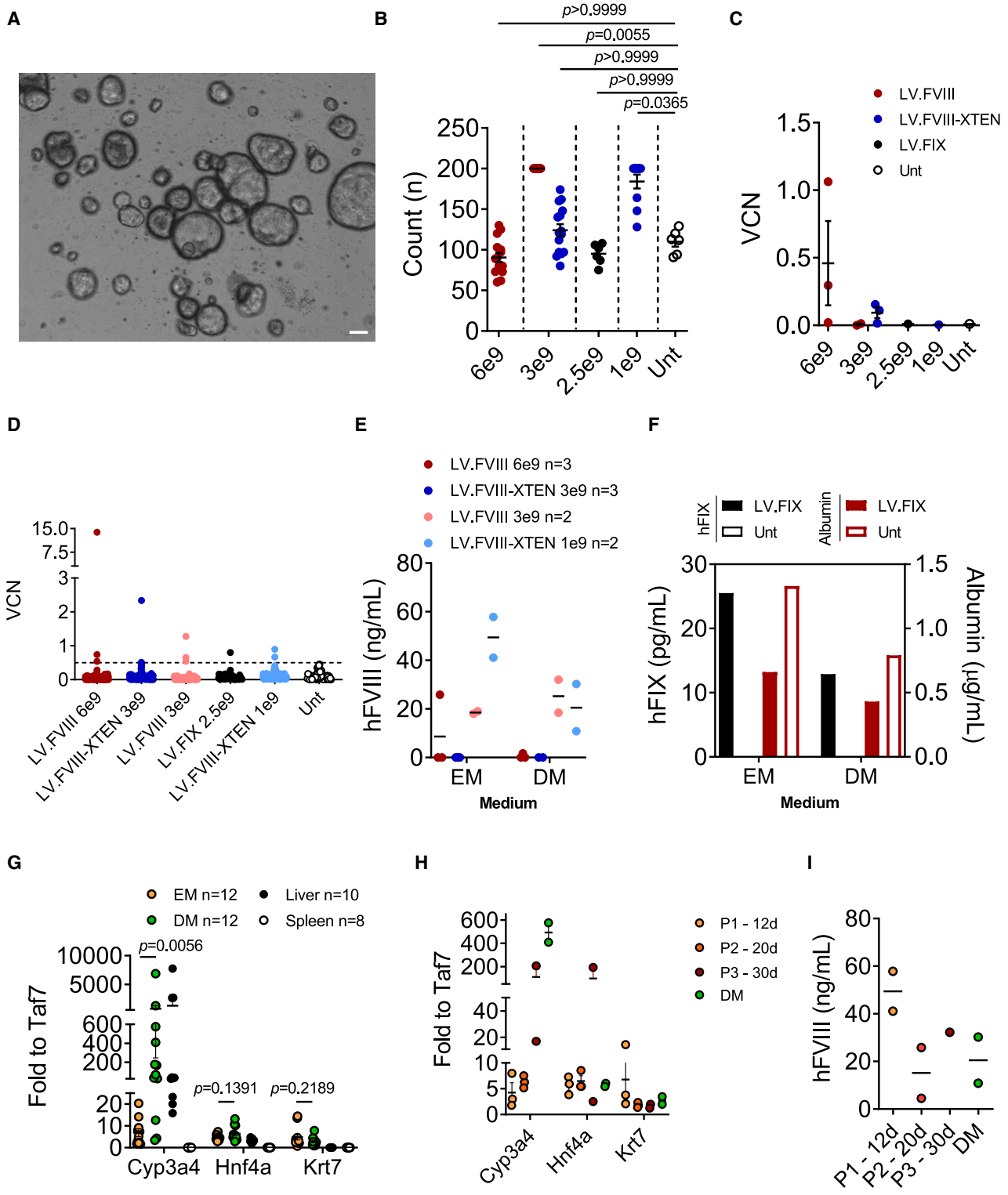


Figure 6. LV gene therapy in NHPs allows targeting cholangiocytes that retain organoid formation potential

(A) Representative picture of NHP-derived liver organoids. Scale bar: 150 μm.

(B) Single values and mean with SEM of organoid count obtained from cholangiocytes isolated from NHP livers treated or not (Unt, $n = 6$ technical replicates: 6 wells of organoids seeded from a single NHP donor) with LVs encoding human FVIII (LV.FVIII; $n = 12–15$ technical replicates: 12–15 wells of organoids seeded from 5 NHP donors), half-life extended hFVIII (LV.FVIII-XTEN; $n = 10–15$ technical replicates: 10–15 wells of organoids seeded from 5 NHP donors), or human FIX

(legend continued on next page)

increase in *Cyp3a4* expression paralleled by a decrease in *Krt7* expression induced by the DM when single organoids were cultured on basal matrix extract (BME)-coated plates (Figure S7A). Interestingly, collagen-coated plates seemed to push toward the cholangiocyte lineage, as shown by the high *Krt7* expression, while culture on plastic supported organoid growth, but gene expression did not significantly change between EM or DM (Figures S7B and S7C). Accordingly, transgene expression, driven by the hepatocyte-specific promoter, was higher in organoids cultured on BME-coated plates compared to the other conditions for both transgenes, and it was undetectable in the organoids derived from the untreated NHP, as expected (Figures S7D–S7F). These data show that NHP organoid differentiation can be polarized by changing the matrix stiffness. Albumin secretion by the organoids did not significantly change between the culture conditions and was independent of LV treatment (Figures S7G and S7H). NHP organoids can be efficiently transduced *ex vivo* by LVs, as shown by VCN and reporter transgene expression in organoids derived from the untreated NHP (Figures S7I and S7J). We could expand organoids from both an untreated and an LV.FIX-transduced NHP enough for an *in vivo* transplantation experiment (Figure 7A). We did not detect substantial differences in the morphology or targeted gene expression between the organoids derived from the treated or untreated NHP (Figures 7B and 7C). We then evaluated the ability of organoids derived from the LV-transduced or untransduced NHP to engraft in immune-deficient mice upon mild liver damage (see Figure 7A). We detected NHP albumin in the blood of the immune-deficient mice transplanted with the liver organoids, indicating engraftment lasting at least 4 months post-administration (Figure 7D). Similar NHP albumin amounts were produced by the organoids derived from an untreated and an LV.FIX-transduced NHP, suggesting equivalent engrafting potential. Indeed, from analyses shown in Figure 6D, we expect that only 2% of the organoids derived from BECs isolated from the LV-treated NHP are actually LV positive; thus, we do not expect differences in engraftment or albumin production compared to organoids derived from BECs isolated from the untreated NHP. We could detect some human FIX in the blood of mice transplanted with the organoids derived from the LV.FIX-transduced NHP, suggesting that *in vivo* LV-transduced cholangiocytes can engraft and express the transgene (Figure 7E). Overall, these data show that NHP cholangiocytes are amenable to *in vivo* LV transduction, express a transthyretin-based promoter, and maintain *in vitro* organoid forming and *in vivo* engrafting potential after gene transfer.

DISCUSSION

We describe here the identification of hepatocyte-primed BECs that are present in the liver in homeostatic conditions without tissue damage, endowed with regenerative potential, and highly permissive to *in vivo* LV-mediated gene transfer. We report that i.v. administration of LVs results in efficient and long-term gene transfer into hepatic BECs in both mice and NHPs. This finding opens the possibility of targeting this cell type both for therapeutic purposes and biotechnological applications. LV-transduced cholangiocytes retain organoid-forming potential, proliferate in response to a toxic insult, and engraft in mice subjected to mild liver damage. The ultimate goal of gene therapy for genetic liver diseases is lifelong reconstitution of the missing function. In order to achieve this goal, the therapeutic transgene needs to persist following homeostatic renewal and potentially damage-induced regeneration of the tissue. Our results represent a favorable outcome toward this goal, as LV-transduced liver cholangiocytes may contribute to transgene expression and replace hepatocytes following tissue damage, thus allowing for maintenance of the therapeutic effect. This is particularly relevant in the challenging setting of inherited metabolic diseases causing liver damage. The lower *in vitro* clonogenic potential and hepatocyte-biased transcriptome displayed by LV-transduced BECs compared to untransduced counterparts did not prevent their long-term maintenance and might even enhance their capacity to respond to liver damage *in vivo*, as shown here in response to cholestatic liver damage. We provide evidence of intrinsic heterogeneity among murine cholangiocytes at a steady state by highlighting a phenotypically defined less-organoid-forming subset, expressing high levels of albumin and HNF1β and characterized by a hepatocyte-biased transcriptome. We captured this until-now elusive cellular state as more permissive to LV gene transfer *in vivo*, suggesting a spatially defined location. We noticed that LV transduction preferentially occurred in cholangiocytes belonging to larger ducts and that genes specific to this cell subset were enriched in transduced vs. untransduced cells. This might be due to their anatomical location and better accessibility to i.v.-injected LVs and/or to different cell-intrinsic permissiveness to LV transduction. Indeed, we found some genes belonging to the Toll-like receptor family and NLRP (nucleotide-binding oligomerization domain, leucine-rich repeat, and Pyrin domain-containing) family overexpressed in the GFP-negative fraction of BECs that may act as restriction factors to LV transduction. We also identified hepatocyte-primed BECs by high albumin expression in an

(LV.FIX, $n = 6$ technical replicates: 6 wells of organoids seeded from a single NHP donor) at the indicated vector doses (TU/kg). For counts higher than 200, 200 has been indicated as the total count. Kruskal-Wallis test with Dunn's multiple comparisons test.
(C) Single values and mean with SEM of VCN of bulk culture of organoids grown from cholangiocytes isolated from NHP treated with the indicated LV at the indicated dose (TU/kg) or left untreated ($n = 1$–3 cultures).
(D) Single values of VCN of single organoids grown from BECs isolated from NHP treated with the indicated LV at the indicated dose (TU/kg) or left untreated ($n = 52$–211 single organoids). Around 2% of analyzed organoids are LV positive (VCN > 0.5, dashed line).
(E and F) Single values and mean of hFVIII (LV.FVIII 6×10^9 $n = 3$, LV.FVIII-XTEN 3×10^9 $n = 3$, LV.FVIII 3×10^9 $n = 2$, and LV.FVIII-XTEN 1×10^9 $n = 2$, E) or hFIX and NHP albumin ($n = 1$, F) measured in the supernatant of organoids cultured in EM or DM derived from LV-treated or untreated NHP, as indicated.
(G) Single values and mean with SEM of the expression of hepatocyte- (*Cyp3a4* and *Hnf4a*) or cholangiocyte- (*Krt7*) specific genes in liver organoids cultured in EM ($n = 12$) or DM ($n = 12$). NHP liver ($n = 10$) and spleen ($n = 8$) samples are shown as controls. Mann-Whitney test.
(H and I) Single values and mean with SEM of the relative expression of the indicated genes (H) or hFVIII concentration measured in the supernatants (I) of organoids cultured in EM overtime, as indicated, or cultured in DM.

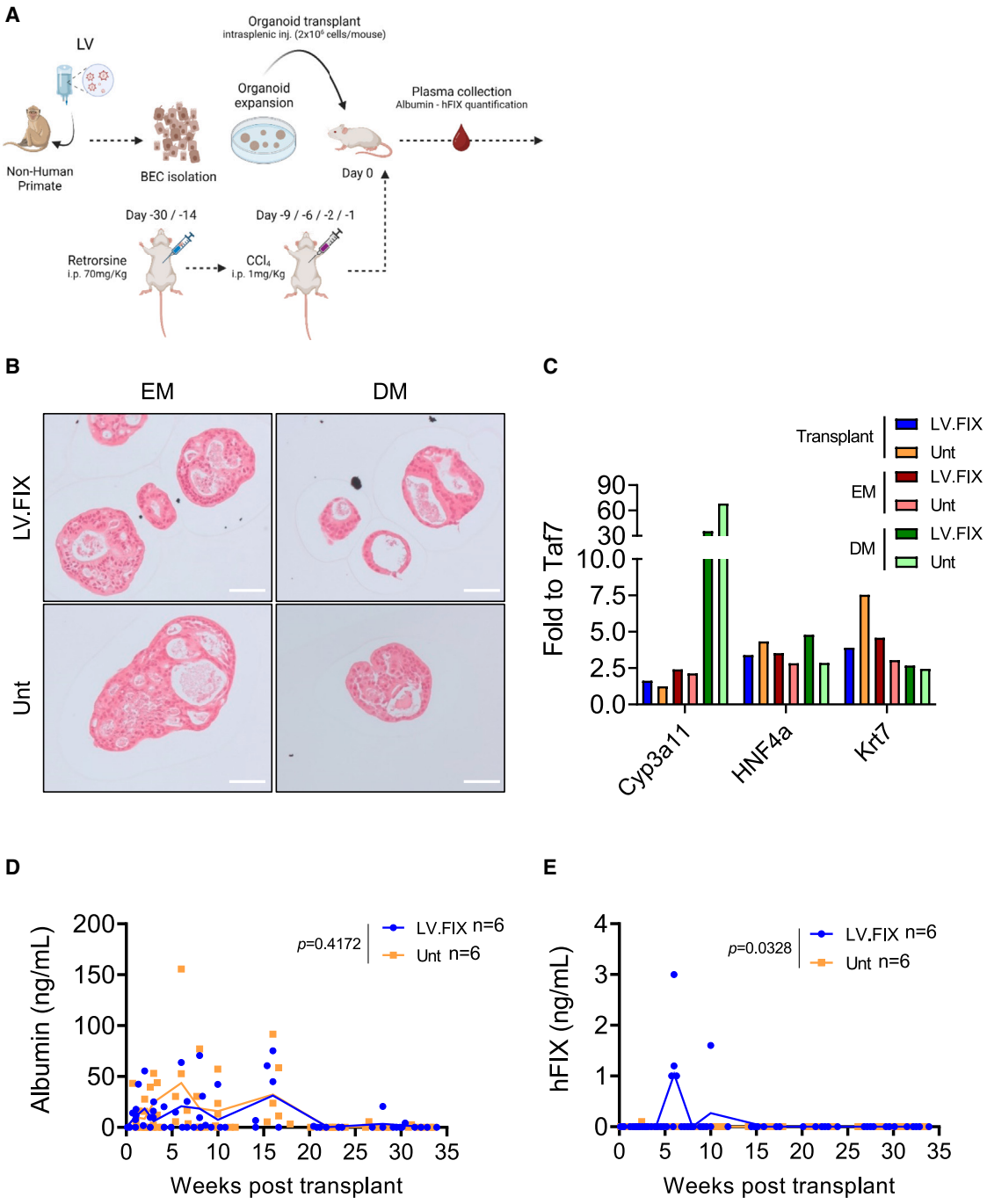


Figure 7. *In vivo* engraftment of organoids derived from LV-treated or untreated NHP

(A) Experimental scheme: BEC are isolated from LV-treated or untreated NHP, and then cultured to obtain liver organoids, which are expanded in EM. NSG mice are pretreated with retrorsine and CCl_4 to allow organoid engraftment. Two million organoids are transplanted by intrasplenic injection. After transplantation, plasma samples are collected from transplanted mice to quantify NHP albumin or hFIX overtime from week 1 to 32 after transplant.

(B) Representative hematoxylin/eosin-stained sections of liver organoids derived from an untreated or an LV.FIX-treated NHP cultured in EM or DM. Scale bar: 100 μm.

(C) Gene expression analysis of the indicated genes in the organoids expanded and transplanted in NSG mice or cultured in EM or DM (n = 1).

(D and E) Single values and mean of NHP albumin (I) or hFIX (J) measured in the plasma of NSG mice transplanted with organoids derived from LV-treated (n = 6, in blue) or untreated (n = 6, in orange) NHP. Two-way ANOVA.

LV-independent manner. Indeed, LV-transduced cholangiocytes have a similar transcriptional profile to those marked by recombination of genetic reporters mediated by Cre expressed under the control of an albumin promoter, further suggesting their hepatocyte priming. In line with these observations, this subset also expressed a hepatocyte-specific transthyretin-based promoter upon *in vivo* LV transduction. Further analyses of this cell subset may unravel specific surface markers amenable to their enrichment or isolation. Liver cholangiocytes isolated from 1-year-old mice were still capable of generating organoids at high efficiency and did not show differences in clonogenic potential whether transduced or not, suggesting progressive enrichment of the more self-renewing compartment over the more committed progenitors *in vivo*. A subset of human cholangiocytes characterized by high TROP2 expression, low organoid formation efficiency, and a hepatocyte-biased transcriptional profile has been previously described.[13] In contrast to the human counterparts, TROP2 expression did not segregate from organoid-forming potential in mice. By comparing the global transcriptional landscape of transduced and untransduced liver cholangiocytes, we found substantial differences, particularly in genes related to hepatocyte-specific processes that were overexpressed in the former compared to the latter. Bioinformatic analysis showed that this transcriptional program was orchestrated by HNF4α, a master regulator of hepatocyte identity.[25] HNF4α controls the proliferation of hepatocytes during fetal development and maintains hepatocyte differentiation in the adult liver.[26] A recent report highlighted the presence of bipotent transitional liver progenitor cells, originating from BECs, that became apparent only after tissue damage and hepatocyte senescence. These transitional liver progenitors co-express HNF4α and the BEC marker cytokeratin 19 before further differentiating into mature hepatocytes.[27] These data are consistent with our identification of a committed BEC state already present in homeostatic conditions that may well represent the population giving rise to these transitional progenitor cells. Our work supports the BEC origin of the bipotent liver facultative stem cells activated following tissue damage. These hepatocyte-primed mouse cholangiocytes can be isolated by exploiting i.v. administration of LV-expressing marker transgenes. Systemic LV delivery also allowed the transduction of NHP cholangiocytes without impacting their capacity to engraft following transplant in immune-deficient mice exposed to mild liver damage. Interestingly, single-cell transcriptomic analysis of NHP livers highlighted a BEC subset enriched for the gene signature identified in *in vivo* LV-transduced murine BECs. Moreover, the activity of the hepatocyte-specific transthyretin promoter in the transduced NHP cholangiocytes further suggests targeting a population similar to that described in mice. In line with this consideration, the bipotent transitional progenitors described by Pu et al. express the hepatocyte-specific transthyretin gene in addition to HNF4α. Heterogeneity among human cholangiocytes has been previously reported, recognizing a less self-renewing and more hepatocyte-biased subset. Reanalysis of the transcriptome of this subset revealed an enrichment of the gene signature identified in the above-mentioned hepatocyte-primed murine BECs. Therefore, our findings related to murine and NHP cholangiocytes are consistent with those regarding human BECs, further supporting their therapeutic relevance, and represent the basis to further investigate the role of these cell subsets in liver regeneration, cholangiocarcinoma formation, and therapeutic applications.

Limitations of the study

The identification of a hepatocyte-primed BEC subset in homeostatic conditions is primarily based on evidence from transcriptomic and *in vitro* analyses. In mouse experiments, the reported preferential proliferation of GFP-positive BECs over the non-transduced ones, together with the hepatocyte-biased gene expression signature of the former, suggests that those cells are a subtype of BECs more prone to respond to damage and potentially *trans*-differentiate into hepatocytes. Additional studies will be required to assess the actual *in vivo trans*-differentiation potential of this BEC subset in conditions supporting robust BEC-to-hepatocyte conversion. Investigation of BEC heterogeneity in NHPs is limited by the impossibility of separating LV-positive and -negative subsets since the LV contained a therapeutic transgene in this study.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Alessio Cantore (cantore.alessio@hsr.it).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- RNA-seq data have been deposited at GEO: GSE232372 and are publicly available as of the date of publication, specifically GEO: GSE232370 (GFP, Figure 3) and GSE232369 (Confetti, Figure 4).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors wish to thank the Flow Cytometry Resource, Advanced Cytometry Technical Applications Laboratory (FRACTAL), and the Advanced Light and Electron Microscopy BioImaging Center (ALEMBIC), core facilities of IRCCS San Raffaele Scientific Institute; the San Raffaele University Center for Statistics in the Biomedical Sciences (CUSSB) for statistical consultancy; the San Raffaele Center for Omics Sciences (COSR); Naldini's and Cantore's laboratories for useful discussions; and Hans Clevers for providing important reagents for organoid culture. This work was mainly supported by the Fondazione Telethon SR-Tiget Core Grant (TTACC0422TT to A.C.) and the EU Horizon 2020 Program (825825 UPGRADE). F. Starinieri conducted this study in partial fulfillment of his international course in molecular medicine at San Raffaele University, Milan.

AUTHOR CONTRIBUTIONS

M.M. designed and performed experiments, analyzed and interpreted data, and wrote the manuscript. F. Starinieri performed experiments and analyzed data. S.B. performed bioinformatic analyses. A.F., T.P., C.C., M.B., F.R., and R.N. provided crucial technical support to experiments. V.B. analyzed immunohistochemistry and immunofluorescence images. F. Sanvito performed histology analyses. I.M. supervised bioinformatic analyses. L.A. and M.H. provided crucial support to organoid cultures and intellectual input. L.N. supervised the work, provided intellectual input, and edited the manuscript.

A.C. supervised and coordinated the work, interpreted data, and wrote the manuscript.

DECLARATION OF INTERESTS

L.N. and A.C. are inventors on patent applications submitted by Foundation Telethon and San Raffaele Scientific Institute on LV technology related to the work presented in this manuscript. M.H. is an inventor of several patents related to organoid technology.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- STUDY DESIGN
- EXPERIMENTAL MODEL AND STUDY PARTICIPANTS
  - Plasmid construction
- METHOD DETAILS
  - Vector production
  - LV titration
  - VCN determination
  - Organoid assays
  - *In vitro* experiments
  - RNA extraction and ddPCR
  - RNA-sequencing
  - Mice experiments
  - Immunofluorescence imaging
  - Immunohistochemistry imaging
  - Histology of organoids
  - ELISA
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.115341.

Received: May 16, 2023
Revised: December 6, 2024
Accepted: January 31, 2025
Published: February 24, 2025

REFERENCES

1. Campana, L., Esser, H., Huch, M., and Forbes, S. (2021). Liver regeneration and inflammation: from fundamental science to clinical applications. Nat. Rev. Mol. Cell Biol. *22*, 608–624. https://doi.org/10.1038/s41580-021-00373-7.
2. Yanger, K., Knigin, D., Zong, Y., Maggs, L., Gu, G., Akiyama, H., Pikarsky, E., and Stanger, B.Z. (2014). Adult hepatocytes are generated by self-duplication rather than stem cell differentiation. Cell Stem Cell *15*, 340–349. https://doi.org/10.1016/j.stem.2014.06.003.
3. Huch, M., and Dollé, L. (2016). The plastic cellular states of liver cells: Are EpCAM and Lgr5 fit for purpose? Hepatology *64*, 652–662. https://doi.org/10.1002/hep.28469.
4. Font-Burgada, J., Shalapour, S., Ramaswamy, S., Hsueh, B., Rossell, D., Umemura, A., Taniguchi, K., Nakagawa, H., Valasek, M.A., Ye, L., et al. (2015). Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer. Cell *162*, 766–779. https://doi.org/10.1016/j.cell.2015.07.026.
5. Munoz-Canoves, P., and Huch, M. (2018). Definitions for adult stem cells debated. Nature *563*, 328–329. https://doi.org/10.1038/d41586-018-07175-6.
6. Huch, M., Dorrell, C., Boj, S.F., van Es, J.H., Li, V.S.W., van de Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M.J., et al. (2013). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature *494*, 247–250. https://doi.org/10.1038/nature11826.
7. Huch, M., Gehart, H., van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M.M.A., Ellis, E., van Wenum, M., Fuchs, S.A., de Ligt, J., et al. (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. Cellule *160*, 299–312. https://doi.org/10.1016/j.cell.2014.11.050.
8. Raven, A., Lu, W.Y., Man, T.Y., Ferreira-Gonzalez, S., O'Duibhir, E., Dwyer, B.J., Thomson, J.P., Meehan, R.R., Bogorad, R., Koteliansky, V., et al. (2017). Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. Nature *547*, 350–354. https://doi.org/10.1038/nature23015.
9. Manco, R., Clerbaux, L.A., Verhulst, S., Bou Nader, M., Sempoux, C., Ambroise, J., Bearzatto, B., Gala, J.L., Horsmans, Y., van Grunsven, L., et al. (2019). Reactive cholangiocytes differentiate into proliferative hepatocytes with efficient DNA repair in mice with chronic liver injury. J. Hepatol. *70*, 1180–1191. https://doi.org/10.1016/j.jhep.2019.02.003.
10. Russell, J.O., Lu, W.Y., Okabe, H., Abrams, M., Oertel, M., Poddar, M., Singh, S., Forbes, S.J., and Monga, S.P. (2019). Hepatocyte-Specific beta-Catenin Deletion During Severe Liver Injury Provokes Cholangiocytes to Differentiate Into Hepatocytes. Hepatology *69*, 742–759. https://doi.org/10.1002/hep.30270.
11. Lu, W.Y., Bird, T.G., Boulter, L., Tsuchiya, A., Cole, A.M., Hay, T., Guest, R.V., Wojtacha, D., Man, T.Y., Mackinnon, A., et al. (2015). Hepatic progenitor cells of biliary origin with liver repopulation capacity. Nat. Cell Biol. *17*, 971–983. https://doi.org/10.1038/ncb3203.
12. Sampaziotis, F., Muraro, D., Tysoe, O.C., Sawiak, S., Beach, T.E., Godfrey, E.M., Upponi, S.S., Brevini, T., Wesley, B.T., Garcia-Bernardo, J., et al. (2021). Cholangiocyte organoids can repair bile ducts after transplantation in the human liver. Science *371*, 839–846. https://doi.org/10.1126/science.aaz6964.
13. Aizarani, N., Saviano, A., Sagar, N., Mailly, L., Durand, S., Herman, J.S., Pessaux, P., Baumert, T.F., and Grün, D. (2019). A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature *572*, 199–204. https://doi.org/10.1038/s41586-019-1373-2.
14. Zabaleta, N., Unzu, C., Weber, N.D., and Gonzalez-Aseguinolaza, G. (2023). Gene therapy for liver diseases - progress and challenges. Nat. Rev. Gastroenterol. Hepatol. *20*, 288–305. https://doi.org/10.1038/s41575-022-00729-0.
15. Ozelo, M.C., Mahlangu, J., Pasi, K.J., Giermasz, A., Leavitt, A.D., Laffan, M., Symington, E., Quon, D.V., Wang, J.D., Peerlinck, K., et al. (2022). Valoctocogene Roxaparvovec Gene Therapy for Hemophilia A. N. Engl. J. Med. *386*, 1013–1025. https://doi.org/10.1056/NEJMoa2113708.
16. Pipe, S.W., Leebeek, F.W.G., Recht, M., Key, N.S., Castaman, G., Miesbach, W., Lattimore, S., Peerlinck, K., Van der Valk, P., Coppens, M., et al. (2023). Gene Therapy with Etranacogene Dezaparvovec for Hemophilia B. N. Engl. J. Med. *388*, 706–718. https://doi.org/10.1056/NEJMoa2211644.
17. Cantore, A., Ranzani, M., Bartholomae, C.C., Volpin, M., Valle, P.D., Sanvito, F., Sergi, L.S., Gallina, P., Benedicenti, F., Bellinger, D., et al. (2015). Liver-directed lentiviral gene therapy in a dog model of hemophilia B. Sci. Transl. Med. *7*, 277ra28. https://doi.org/10.1126/scitranslmed.aaa1405.
18. Milani, M., Annoni, A., Moalli, F., Liu, T., Cesana, D., Calabria, A., Bartolaccini, S., Biffi, M., Russo, F., Visigalli, I., et al. (2019). Phagocytosis-shielded lentiviral vectors improve liver gene therapy in nonhuman primates. Sci. Transl. Med. *11*, eaav7325. https://doi.org/10.1126/scitranslmed.aav7325.
19. Milani, M., Canepari, C., Liu, T., Biffi, M., Russo, F., Plati, T., Curto, R., Patarroyo-White, S., Drager, D., Visigalli, I., et al. (2022). Liver-directed lentiviral gene therapy corrects hemophilia A mice and achieves normal-range factor VIII activity in non-human primates. Nat. Commun. *13*, 2454. https://doi.org/10.1038/s41467-022-30102-3.

20. Milani, M., Canepari, C., Assanelli, S., Merlin, S., Borroni, E., Starinieri, F., Biffi, M., Russo, F., Fabiano, A., Zambroni, D., et al. (2024). GP64-pseudotyped lentiviral vectors target liver endothelial cells and correct hemophilia A mice. EMBO Mol. Med. *16*, 1427–1450. https://doi.org/10.1038/s44321-024-00072-8.

21. Tabibian, J.H., Masyuk, A.I., Masyuk, T.V., O'Hara, S.P., and LaRusso, N.F. (2013). Physiology of cholangiocytes. Compr. Physiol. *3*, 541–565. https://doi.org/10.1002/cphy.c120019.

22. Isse, K., Lesniak, A., Grama, K., Maier, J., Specht, S., Castillo-Rama, M., Lunz, J., Roysam, B., Michalopoulos, G., and Demetris, A.J. (2013). Preexisting epithelial diversity in normal human livers: a tissue-tethered cytometric analysis in portal/periportal epithelial cells. Hepatology *57*, 1632–1643. https://doi.org/10.1002/hep.26131.

23. Snippert, H.J., Haegebarth, A., Kasper, M., Jaks, V., van Es, J.H., Barker, N., van de Wetering, M., van den Born, M., Begthel, H., Vries, R.G., et al. (2010). Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. Science *327*, 1385–1389. https://doi.org/10.1126/science.1184733.

24. Broutier, L., Andersson-Rolf, A., Hindley, C.J., Boj, S.F., Clevers, H., Koo, B.K., and Huch, M. (2016). Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. Nat. Protoc. *11*, 1724–1743. https://doi.org/10.1038/nprot.2016.097.

25. Dubois, V., Staels, B., Lefebvre, P., Verzi, M.P., and Eeckhoute, J. (2020). Control of Cell Identity by the Nuclear Receptor HNF4 in Organ Pathophysiology. Cell *9*, 2185. https://doi.org/10.3390/cells9102185.

26. Bonzo, J.A., Ferry, C.H., Matsubara, T., Kim, J.H., and Gonzalez, F.J. (2012). Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice. J. Biol. Chem. *287*, 7345–7356. https://doi.org/10.1074/jbc.M111.334599.

27. Pu, W., Zhu, H., Zhang, M., Pikiolek, M., Ercan, C., Li, J., Huang, X., Han, X., Zhang, Z., Lv, Z., et al. (2023). Bipotent transitional liver progenitor cells contribute to liver regeneration. Nat. Genet. *55*, 651–664. https://doi.org/10.1038/s41588-023-01335-9.

28. Mátrai, J., Cantore, A., Bartholomae, C.C., Annoni, A., Wang, W., Acosta-Sanchez, A., Samara-Kuko, E., De Waele, L., Ma, L., Genovese, P., et al. (2011). Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. Hepatology *53*, 1696–1707. https://doi.org/10.1002/hep.24230.

29. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

30. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

31. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L., Fang, T., Doncheva, N.T., Pyysalo, S., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.

32. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.

33. Schuler, M., Dierich, A., Chambon, P., and Metzger, D. (2004). Efficient temporally controlled targeted somatic mutagenesis in hepatocytes of the mouse. Genesis *39*, 167–172. https://doi.org/10.1002/gene.20039.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
APC-conjugated anti-mouse EpCAM	Invitrogen	Clone: G8.8; Cat#17-5791-82
PE-*anti*-mouse CD31	BD Pharmigen	Clone: MEC 13.3; Cat#561073
PE-*anti*-mouse CD45	BioLegend	Clone: 104; Cat#109807
Rabbit anti-GFP	Invitrogen	Clone: polyclonal; Cat#A11122
anti-mouse HNF4α	Cell Signaling Technologies	Clone: C11F12; Cat#3113S
anti-mouse HNF1β	Proteintech	Clone: polyclonal; Cat#12533-1-AP
anti-hFVIII binding Ab	Green Mountain Antibodies	Cat#GMA8016
hFVIII detection Ab for ELISA	Affinity Biologicals	Cat#F8C-EIC-D
anti-hFIX binding Ab	Haematologic Technologies	Cat#AHIX-5041
hFIX detection Ab for ELISA	Cederlane	Cat#CL20040APHP
Rabbit anti-mouse CK7	Abcam	Clone: EPR17078; Cat#ab181598
Chicken anti-GFP	Invitrogen	Clone: polyclonal; Cat#A10262
APC-*anti*-NGFR	Miltenyi	Clone: ME20.4–1.H4; Cat#130-113-418
Donkey anti-rabbit IgG-AF647	Invitrogen	Cat#A31573
Goat anti-chicken IgY-AF488	Invitrogen	Cat#A11039
Biological samples		
Non-human primate liver samples	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Matrigel matrix	BD biosciences	Cat#356231
Basal membrane extract (BME)	Cultrex	Cat# 3533-005-02
Tamoxifen	Sigma-Aldrich	Cat#T5648
3,5-Diethoxycarbonyl-1,4-Dihydrocollidine (DDC)	Sigma-Aldrich	Cat#137030-25G
Retrorsine	Sigma-Aldrich	Cat#R0382-100MG
ReFACTO	Pfizer	Cat#034421038E
BeneFIX 250 UI/mL	Pfizer	N/A
B27	ThermoFisher	Cat#12587-010
N-acetylcysteine	Sigma-Aldrich	Cat#A8199
Rspo1-conditioned medium	This paper	N/A
Nicotinamide	Sigma-Aldrich	Cat#N0636
Rh[leu15]-gastrin I	Sigma-Aldrich	Cat#G9145
rmEGF	R&D	Cat#2028-EG-200
rhFGF10	R&D	Cat#345-FG-025
rhHGF	R&D	Cat#394-HG-025
hNoggin	R&D	Cat#6057-NG-025
Wnt3a-conditioned medium	This paper	N/A
ROCK inhibitor Y-27632	Sigma-Aldrich	Cat#Y0503
A83-01	Tocris Bioscience	Cat#2939
DAPT	Sigma-Aldrich	Cat#D5942
Dexamethasone	Sigma-Aldrich	Cat#D4902
N2 supplement	TermoFischer	Cat#17502-048
rhEGF	R&D	Cat#236-EG-200
Forskolin	Tocris Bioscience	Cat#1099/10
B27 with vitamin A	TermoFisher	Cat#17504044

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BMP7	R&D	Cat#354-BP-010
rhFGF19	R&D	Cat#969-FG-025
Critical commercial assays		
HIV-1 Gag p24 antigen immunocapture assay	Perkin Elmer	Cat#NEK050B
MagMAX mirVana total RNA	ThermoFisher	Cat#A27828
SuperScript IV VILO kit	ThermoFisher	Cat#11766050
SMART-Seq v4 for mRNA-seq kit	Takara	Cat#112219
Bond Polymer Refine Detection	Leica	Cat#DS9800
Mouse albumin ELISA Kit	Abcam	Cat#AB108792
AssayMax monkey albumin ELISA Kit	AssayPro	Cat#EKA3201-1
Deposited data		
Human cell ATLAS	Aizarani et al.[13]	GEO Project code: GSE124395
Raw and analyzed data	This paper	GEO Project code: GSE232372
Experimental models: Cell lines		
Human: Hek293T	Luigi Naldini's Lab	N/A
Experimental models: Organisms/strains		
Mouse: C57BL/6	Charles River Laboratories	stock #000664
Mouse: NSG	Charles River Laboratories	stock #005557
Mouse: B6.129P2-Gt(ROSA)26Sortm1 (CAG-Brainbow2.1)Cle/J	The Jackson Laboratories	stock #017492
Mouse: $SA^{+/CreERT2}$	Matteo Iannacone's Lab	Schuler et al.[29]
Oligonucleotides		
HIV fw: 5′-T ACTGACGCTCTCGCACC-3'; HIV rv: 5′-TCTCGACGCAGGACTCG-3'; HIV Probe FAM 5′-ATCTCTCTCCTTCTAGCCTC-3′	Matrai et al.[28]	N/A
Telo fw: 5′-GGCACACGTGGCTTTTCG-3'; Telo rv: 5′-GGTGAACCTCGTAAGTTTATGCAA-3'; Telo probe: VIC 5′-TCAGGACGTCGAGTGGACACGGTG-3′ TAMRA	Matrai et al.[28]	N/A
human GAPDH gene	Applied Biosystems	Cat#HS00483111_cm
RT-LV; ΔU3 fw: 5′-TCACTCCCAACGAAGACAAGATC-3′, gag rv: 5′-GAGTCCTGCGTCGAGAGAG-3′	Matrai et al.[28]	N/A
Sema3A fw: 5′-ACCGATTCCAGATGATTGGC-3'; Sema3A rv: 5′-TCCATATTAATGCAGTGCTTGC-3'; Sema3A probe: HEX 5′-AGAGGCCTGTCCTGC AGCTCATGG-3′ BHQ1	Milani et al.[18]	N/A
NHP TATA-Box Binding Protein Associated Factor 7 (TAF7) gene	Applied Biosystems	Cat#RH 02916247_s1
Murine HPRT	Biorad	Cat#dMmuCPE5095493
See "RNA extraction and ddPCR" section for additional gene expression primer set	This paper	N/A
Recombinant DNA		
pCCLsin.cPPT.ET.GFP.142T	Matrai et al.[28]	N/A
pCCLsin.cPPT.PGK.GFP	Matrai et al.[28]	N/A
pMDLg/pRRE	Milani et al.[18]	N/A
pCMV.REV	Milani et al.[18]	N/A
pMD2.G	Milani et al.[18]	N/A
pAdVantage	Promega	Cat#E1711

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
DIVA software	BD Biosciences	N/A
QuantaSoft software	BioRad	N/A
R/Bioconductor package DESeq2	Love et al.[29]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
R/Bioconductor package ClusterProfiler on KEGG Pathways, Panglao database, and ChEA transcription factors database	Wu et al.[30]	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
StringDB	Szklarczyk et al.[31]	https://string-db.org/
CellRanger software (v7.0.1)	10X Genomics	https://www.10xgenomics.com/support/software/cell-ranger/
R environment (v.4.3.2)	N/A	https://cran.r-project.org/
Seurat (v5.1.0)	Hao et al.[32]	https://cloud.r-project.org/web/packages/Seurat/index.html
ImageJ	https://imagej.nih.gov/ij/	N/A
ArivisVision4D	ZeissAG	N/A
Imagescope	Leica Biosystem	N/A

STUDY DESIGN

Sample size in experiments with mice was chosen according to previous experience with experimental models and assays. No sample or animal was excluded from the analyses. Mice were randomly assigned to each experimental group. Investigators were not blinded.

EXPERIMENTAL MODEL AND STUDY PARTICIPANTS

The human embryonic kidney (HEK)-293T cell line used to produce LV was maintained in Iscove's modified Dulbecco's medium (IMDM, Sigma) supplemented with 10% fetal bovine serum (FBS, FetalClone II, HyClone, Euroclone), 4 mM glutamine (Lonza), penicillin and streptomycin 100 international units (IU)/mL (Lonza). Cells were maintained in a 5% CO_2 humidified atmosphere at 37°C. This cell line was routinely tested for mycoplasma contamination and was not authenticated. Primary BEC were isolated from mice or NHP (see "organoid assays" below).

Male and female C57BL/6 and NSG mice were purchased from Charles River Laboratories. Founder B6.129P2-Gt(ROSA)26Sortm1(CAG-Brainbow2.1)Cle/J mice (referred to as Confetti mice[23]) were obtained from The Jackson Laboratories (stock #017492). Founder $SA^{+/CreERT2}$ mice (referred to as Alb-CreERT mice[33]) were obtained from Matteo Iannacone's Lab (IRCCS Ospedale San Raffaele). All mice were maintained in specific pathogen-free conditions and fed *ad libitum*. Alb-CreERT/Confetti mice were obtained by crossing Confetti homozygous mice with Alb-CreERT homozygous mice. All animal procedures were performed according to protocols approved by the Institutional Animal Care and Use Committee.

Plasmid construction

Plasmid pCCLsin.cPPT.ET.GFP.142T and pCCLsin.cPPT.PGK.GFP were previously described.[28]

METHOD DETAILS

Vector production

Third-generation self-inactivating (SIN) LV were produced by calcium phosphate transient transfection into 293T cells. 293T cells were transfected with a solution containing a mix of the selected LV genome transfer plasmid, the packaging plasmids pMDLg/pRRE and pCMV.REV, pMD2.G and pAdVantage (Promega), as previously described.[18] Briefly, medium was changed 14–16 h after transfection and supernatant was collected 30 h after medium change. LV-containing supernatants were sterilized through a 0.22μm filter (Millipore), transferred into sterile poliallomer tubes (Beckman), and centrifuged at 20,000 g for 120 min at 20°C (Beckman Optima XL-100K Ultracentrifuge). LV pellet was dissolved in the appropriate volume of PBS to allow 500X concentration. Bald.LV were produced in absence of envelope-encoding packaging plasmid, while Empty.LV were produced in absence of the LV genome-encoding plasmid.

LV titration

For LV titration, 1x10^5 293T cells were transduced with serial LV dilutions in the presence of polybrene (8 μg/mL). For LV-PGK.GFP, cells were analyzed by flow cytometry using a FACSCanto analyzer (BD Biosciences), equipped with DIVA Software, 5–7 days after transduction, and infectious titer, expressed as transducing units (TU)/mL, was calculated using the formula TU/mL = ((% GFP+ cells/100)x100,000x(1/dilution factor)). For all other LV, genomic DNA Genomic DNA (gDNA) was extracted 14 days after transduction, using Maxwell 16 Cell DNA Purification Kit (Promega), following manufacturer's instructions. VCN was determined was determined by ddPCR, starting from 5 to 20 ng of template gDNA using primers (HIV fw: 5′-T ACTGACGCTCTCGCACC-3'; HIV rv: 5′-TCTCG ACGCAGGACTCG-3′) and a probe (FAM 5′-ATCTCTCTCCTTCTAGCCTC-3′) designed on the primer binding site region of LV. The amount of endogenous DNA was quantified by a primers/probe set designed on the human telomerase gene (Telo fw: 5′-GGCAC ACGTGGCTTTTCG-3'; Telo rv: 5′-GGTGAACCTCGTAAGTTTATGCAA-3'; Telo probe: VIC 5′-TCAGGACGTCGAGTGGACACGG TG-3′ TAMRA) or the human GAPDH gene (Applied Biosystems HS00483111_cm). The PCR reaction was performed with each primer (900 nM) and the probe (250 nM, 500nM for Telo) following manufacturer's instructions (Biorad), read with QX200 reader and analyzed with QuantaSoft software (Biorad). Infectious titer, expressed as TU/mL, was calculated using the formula TU/mL = (VCNx100,000x(1/dilution factor). LV physical particles were measured by HIV-1 Gag p24 antigen immunocapture assay (PerkinElmer) following manufacturer's instructions. LV-specific infectivity was calculated as the ratio between infectious titer and physical particles.

VCN determination

For mice experiments, DNA was extracted from whole liver samples using Maxwell 16 Tissue DNA Purification Kit (Promega), while from EpCAM-positive sorted cells or liver organoid samples DNA was extracted using QIAamp DNA Micro Kit (Qiagen), according to manufacturer's instructions. VCN in murine DNA was determined by ddPCR, starting from 5 to 20 ng of template gDNA using a primers/probe set designed on the primer binding site region of LV (see "LV titration" above) or using an ad hoc ddPCR (QX200 EvaGreen Digital PCR Supermix, Bio-Rad), which selectively amplifies the reverse transcribed vector genome (both integrated and non-integrated) discriminating it from plasmid carried over from the transient transfection (RT-LV; ΔU3 fw: 5′-TCACTCCCAACGAAGACAAGATC-3′, gag rv: 5′-GAGTCCTGCGTCGAGAGAG-3′) (Matrai et al., 2011). The amount of endogenous DNA was by a primers/probe set designed on the murine sema3a gene (Sema3A fw: 5′-ACCGATTCCAGATGATTGGC-3'; Sema3A rv: 5′-TCCATATTAATGCAGTGCTTGC-3'; Sema3A probe: HEX 5′-AGAGGCCTGTCCTGCAGCTCATGG-3′ BHQ1). The PCR reaction was performed with each primer (900 nM, 150nM for RT-LV primers) and the probe (250 nM) following manufacturer's instructions (Biorad), read with QX200 reader and analyzed with QuantaSoft software (Biorad). For NHP samples, DNA was extracted using QIAamp DNA Micro Kit (Qiagen), according to manufacturer's instructions, or samples were lysed using QuickExtract DNA extraction solution (Biosearch Technologies) for single organoid analyses. VCN in NHP DNA was determined by ddPCR, starting from 5 to 20 ng of template gDNA using a primers/probe set designed on the primer binding site region of LV (see "LV titration" above) and genomic DNA was quantified by primers set designed on the NHP TATA-Box Binding Protein Associated Factor 7 (TAF7) gene, (Applied Biosystems RH 02916247_s1).

Organoid assays

Single-cell suspension of murine liver ductal cells was obtained following Broutier et al.[24] Briefly, livers were chopped in small pieces and digested for 2 h at 37°C on the shaker with 125–150 μg/mL collagenase/dispase solution, according to liver dimension, replacing it with fresh solution every hour. To obtain single-cell suspension of ductal cells, an additional digestion of 10 min at 37°C was performed using 10x TrypLE solution (Thermo Fisher Scientific). Cells were stained with allophycocyanin (APC)-conjugated anti-EpCAM (G8.8, Invitrogen) and phycoerythrin (PE)-conjugated anti-nPC mix (anti-CD31, MEC 13.3, BD Pharmigen; anti-CD45, 104, BioLegend) and then EpCAM-positive GFP-positive/-negative cells were sorted using BD FACSAria Fusion Cell Sorter (BD Biosciences) or MoFlo Astrios EQ Cell Sorter (Beckman Coulter) and plated in Matrigel matrix (BD biosciences) up to 3,000 cells/40 μL drop, according to number of cells obtained after sorting, in isolation medium.[24] 5 days after the seeding, isolation medium was replaced with expansion medium,[24] and, after additional 7 days, organoids were manually counted using EVOS (Thermo Fisher Scientific) and collected for VCN measurement and gene expression analyses (referred to as P0). One-third of organoids was then split and seeded again in Matrigel matrix in expansion or differentiation medium for 10 additional days,[24] then organoids were counted and collected for VCN measurement and gene expression analyses (referred to as passage 1, P1, or 2, P2 in case organoids were split again during this 10 additional days of culture). Organoid formation efficiency is determined as number of organoids*100/number of seeded cells. Please find below details about isolation or expansion medium. To evaluate organoid formation efficiency of EpCAM-positive Confetti-positive/-negative cells, cholangiocytes were FACS sorted pooling together RFP-, CFP-, and YFP/GFP-positive cells, referred to as "Confetti-positive", or RFP-, CFP-, and YFP/GFP-negative cells, referred to as "Confetti-negative" cells. Single cell suspension of NHP liver ductal cells was obtained following the protocol used for human liver cholangiocytes described in Broutier et al.[24] Briefly, livers were chopped in small pieces and digested for 2 h at 37°C on the shaker with 2.5 mg/mL collagenase/dispase solution (5 mL/g), replacing it with fresh solution every hour. Cells were plated in basal membrane extract (BME, Cultrex) up to 10,000 cells/50 μL drop in isolation medium.[24] NHP-derived organoids were then cultured and analyzed similarly to mouse-derived organoids with the appropriate expansion or differentiation medium.[24] For single organoid culture, organoids were picked and cultured in 96-well plates pre-coated with BME (diluted 1:4-1:10), collagen (Sigma-Aldrich), or left uncoated (plastic). Supernatants for the determination of the concentration of NHP albumin, hFIX or hFVIII (see "ELISA assays") were collected at the end of the culture

at P0 or P1/P2. Expansion medium for murine organoids is prepared by adding the following reagents to basal medium (Advanced DMEM/F12, ThermoFisher; penicillin/streptomycin 100 IU/mL, Lonza; GlutaMax, ThermoFisher; HEPES 10mM, ThermoFisher).

Reagent	Concentration	Vendor
B27	1x	ThermoFisher
N-acetylcysteine	1mM	Sigma-Aldrich
Rspo1-CM*	0.05x	Prepared in house
Nicotinamide	10mM	Sigma-Aldrich
Rh[leu15]-gastrin I	10nM	Sigma-Aldrich
rmEGF	50 ng/mL	R&D
rhFGF10	100 ng/mL	R&D
rhHGF	50 ng/mL	R&D

*CM: conditioned medium.

Isolation medium for murine or NHP organoids is prepared by adding the following reagents to expansion medium.

Reagent	Concentration	Vendor
hNoggin	25 ng/mL	R&D
Wnt3a-CM*	0.3x	Prepared in house
ROCK inhibitor Y-27632	10uM	Sigma-Aldrich

*CM: conditioned medium.

Differentiation medium for murine organoids is prepared by adding the following reagents to basal medium.

Reagent	Concentration	Vendor
B27	1x	TermoFisher
N-acetylcysteine	1mM	Sigma-Aldrich
Rh[leu15]-gastrin I	10nM	Sigma-Aldrich
rmEGF	50 ng/mL	R&D
rhFGF10	100 ng/mL	R&D
A83-01	50nM	Tocris Bioscience
DAPT	10uM	Sigma-Aldrich
Dexamethasone	3uM	Sigma-Aldrich

Expansion medium for NHP organoids is prepared by adding the following reagents to basal medium.

Reagent	Concentration	Vendor
B27	1x	TermoFisher
N2 supplement	1x	TermoFischer
N-acetylcysteine	1mM	Sigma-Aldrich
Rspo1-CM*	0.1x	Prepared in house
Nicotinamide	10mM	Sigma-Aldrich
Rh[leu15]-gastrin I	10nM	Sigma-Aldrich
rhEGF	50 ng/mL	R&D
rhFGF10	100 ng/mL	R&D
rhHGF	25 ng/mL	R&D
Forskolin	10uM	Tocris Bioscience
A83-01	5uM	Tocris Bioscience

*CM: conditioned medium.

Differentiation medium for NHP organoids is prepared by adding the following reagents to basal medium.

Reagent	Concentration	Vendor
B27 with vitamin A	1x	TermoFisher
N2 supplement	1x	TermoFisher
N-acetylcysteine	1mM	Sigma-Aldrich
Rh[leu15]-gastrin I	10nM	Sigma-Aldrich
rhEGF	50 ng/mL	R&D
rhHGF	25 ng/mL	R&D
A83-01	0.5uM	Tocris Bioscience
DAPT	10uM	Sigma-Aldrich
Dexamethasone	3uM	Sigma-Aldrich
BMP7	25 ng/mL	R&D
rhFGF19	10 ng/mL	R&D

In vitro experiments

For *in vitro* transduction experiments, 6,000 EpCAM-positive sorted cells or organoids were transduced by spinoculation (2,400 rpm for 1 h at 32°C) with LV encoding for GFP (PGK.GFP LV), or control Empty.LV or Bald.LV at multiplicity of infection 10.

RNA extraction and ddPCR

RNA extraction was performed using KingFisher protocol MagMAX mirVana total RNA (ThermoFisher Scientific, A27828), according to manufacturer's instructions and reverse transcribed using the SuperScript IV VILO kit (11766050; ThermoFisher Scientific). We stored samples in 350μL of RLTplus (Qiagen) at −80°C, thus we skipped the addition of Lysis buffer (step 1 of the RNA extraction protocol) and we added 350μL of isopropanol (no 2-Mercaptoethanol). LV gene expression was assessed by ddPCR starting from 10 to 25 ng of template cDNA using a primers/probe set designed on the genes listed below. *HPRT* was used as reference gene for mouse samples (dMmuCPE5095493, Biorad), while *Taf7* was used as reference gene for NHP samples (see "VCN determination"). The PCR reaction was performed with each primer (900 nM) and the probe (250 nM) following manufacturer's instructions (Biorad), read with QX200 reader and analyzed with QuantaSoft software (Biorad).

Gene	Species	Vendor	Cat. Number
Alb	Mus musculus	ThermoFisher Scientific	Mm00802090
Hnf4a	Mus musculus	ThermoFisher Scientific	Mm01247712
Hnf1b	Mus musculus	ThermoFisher Scientific	Mm00447459
Cyp3a11	Mus musculus	ThermoFisher Scientific	Mm00731567
EpCAM	Mus musculus	Biorad	dMmu CPE5102938
Krt7	Mus musculus	ThermoFisher Scientific	Mm00466676
Krt17	Mus musculus	ThermoFisher Scientific	Mm00495207
Cyp3a4	Macaca Mulatta	ThermoFisher Scientific	Rh02872540
Hnf4a	Macaca Mulatta	ThermoFisher Scientific	Rh02828709
Krt7	Macaca Mulatta	ThermoFisher Scientific	Rh04253299

RNA-sequencing

For bulk transcriptomic analysis RNA was extracted using RNeasy Micro Kit (Qiagen), following manufacturer's instructions. Library preparation was performed by SMART-Seq v4 for mRNA-seq kit (TaKaRa), then Illumina sequencing was performed at the San Raffaele Center for Omics Sciences or at Azenta Lifesciences. The initial processing of input sequences was performed with FastQC to assess read quality and trimmomatic to get rid of low-quality sequences. Reads were then aligned to the mouse genome assembly (GRCm39) using the STAR software with standard parameters, and abundancies at gene level were computed using the Subread featureCounts function. Differential Gene Expression (DGE) analysis on GFPpos compared to GFPneg cells was done using the R/Bioconductor package DESeq2,[29] normalizing for library size using DESeq2's median of ratios. *p*-values were corrected using false discovery rate (FDR), and genes having FDR <0.01 and abs(logFC) > 2 were considered as differentially expressed. Post-analyses on DGE results were performed with the R/Bioconductor package ClusterProfiler on KEGG Pathways, Panglao database,

and ChEA transcription factors database.[30] StringDB[31] was employed to perform gene network analysis on upregulated DEGs in GFPpos compared to GFPneg cells. For single-cell RNA sequencing (scRNAseq) of NHP BEC, frozen suspensions of cells enriched in cholangiocytes obtained from the enzymatic digestion of liver pieces of 2 NHP (one untreated and one LV-treated at 3e9 TU/kg) were shipped for processing to Azenta. Cells were thawed and underwent two rounds of dead cell removal before the preparation of the cDNA library using 10X Genomics Chromium 3′ gene expression RNA-seq followed by Illumina sequencing. Resulting 2x150 bp reads were aligned against the Macaca Nemestrina (Mnem) reference genome and quantified using the CellRanger software (v7.0.1) with default settings to produce a cell-by-gene matrix for all the samples, which were then imported in the R environment (v.4.3.2) and analyzed with Seurat (v5.1.0).[32] In detail, cells expressing less than 500 (indication of low-viability) or more than 5,000 (indication of potential doublets) genes were excluded. Samples were then merged, and the resulting combined object was analyzed by performing log-normalization with a scale factor of 1,000 by using the *NormalizedData* function of the Seurat package, principal component analysis, batch removal (using Harmony), clustering, and Uniform Manifold Approximation and Projection (UMAP) embedding computation. Cluster markers were computed using the *FindAllMarkers* function of the Seurat package, which exploits a Wilcoxon Rank-Sum test for significance. The Seurat function *AddModuleScore* was used to compute the average expression of gene signatures in the analyzed cells. Clusters from the "Human Cell Atlas" study (GSE124395)[13] were mapped to the dataset using the *FindTransferAnchors* + *TransferData* functions from Seurat.

Mice experiments

Tamoxifen (Sigma-Aldrich T5648) was dissolved in corn oil 10 mg/mL, stored at −20°C and administered subcutaneously in newborn mice at the dose of 0.1 mg/g. Vector administration was carried out in male and female newborns (postnatal day 0 or 1) by temporal vein injection (25–30 μL/mouse) of $5x10^{10}$ TU/kg LV dose. For evaluation of proliferation of LV-transduced or untransduced BEC, LV-treated mice were challenged for 1 week with a diet supplemented with 0.1% of 3,5-Diethoxycarbonyl-1,4-Dihydrocollidine (DDC). Weight of the mice was checked daily during the DDC treatment. For transplantation experiments mice were pre-treated with 2 intraperitoneal (i.p.) injections of retrorsine (70 mg/kg, Sigma-Aldrich) 4 and 2 weeks before transplant to limit hepatocyte-mediated liver regeneration. Retrorsine is resuspended in distilled H_2O at 10 mg/mL, titrated at pH 2.5 to be dissolved completely with HCl 1 mol/L then neutralized to pH 7.0 with NaOH 1 mol/L; final concentration of administered retrorsine is 5 mg/mL diluted right before administration 1:2 with NaCl 0.15 mol/L pH 7.0. To allow for organoid engraftment, mice were subjected to mild liver damage with 4x i.p. administrations of CCl_4 (1 mg/kg) at days 9, 6, 2 and 1 before transplant. For transplantation, mice were anesthetized with isoflurane, subjected to surgery to expose the spleen and $2x10^6$ organoids were injected intrasplenically. Mice were medicated the day of the transplant with Carprofene (5 mg/kg subcutaneously, Rymadil, Pfizer) and with gentamicin (80 mg/mL, Italfarmaco) in drinking water for 1 week after transplant (final concentration 0.3 mg/mL). Mice were bled from the retro-orbital plexus using capillary tubes and blood was collected into 0.38% sodium citrate buffer, pH 7.4 for determination of the concentration of NHP albumin or hFIX (see "ELISA assays"). Mice were bled from the retro-orbital plexus using capillary tubes and blood was collected for determination of the concentration of ALT, AST and ALP. In particular, ALT/glutamic pyruvic transaminase (GTP) liquid (#0018257440); Alkaline Phosphatase (#0018259740); AST/GOT Aspartate aminotransferase/Glutamic oxacetic transaminase Liquid (#0018257540) were used for the quantitative determination of the serum with an International Federation of Clinical Chemistry and Laboratory Medicine-optimized kinetic ultraviolet (UV) method in an ILab650 chemical analyzer (Instrumentation Laboratory). SeraChem Control Level 1 and Level2 (#0018162412 and #0018162512) were analyzed as quality controls. Mice were euthanized by CO_2 inhalation at the scheduled times. All animal procedures were performed according to protocols approved by the Institutional Animal Care and Use Committee.

Immunofluorescence imaging

Livers were harvested from mice, washed briefly in PBS and fixed 4 h in PBS 4% paraformaldehyde (PFA), then washed again briefly in PBS before being stored at least 24 h in 30% sucrose 0.02% sodium azide in H_2O. Livers or murine oprganoids were then frozen in OCT (optimal cutting temperature) compound (Killik, Bio Optica) and slices 5 to 20 μm thick were cut at cryostat (Histo-line MC5050), placed on Superfrost Plus microscope slides (Thermo Scientific) and stored at −80°C. All the staining steps were performed protected from lights to preserve endogenous fluorescence. Slides were thawed at room temperature for at least 2 h, then washed 3 times 5 min in PBS 0.1% X-Triton. Edges were drawn around the tissue using immunostaining pap pen (Sigma-Aldrich) to contain staining solutions. Blocking step was performed using PBS 0.1% X-Triton 1% bovine serum albumin (BSA) 5% FBS for 1 h at room temperature in a humid chamber. Blocking solution was then substitute with mix containing primary antibodies in blocking solution and incubation was carried on for 12–16 h at 4°C in a humid chamber. Primary antibodies mix solution was removed and slides were washed 3 times 5 min in PBS 0.1% X-Triton. Secondary antibodies were diluted in blocking solution together with Hoechst (Invitrogen, 1:20000) and incubated for 1 h at room temperature in a humid chamber. Slides were then washed in PBS and coverslip was added using Fluoromount-G (Invitrogen) as mounting medium. Slides were dried for 16 h at room temperature or for longer at 4°C, and stored at 4°C before acquisition. The following antibodies were used for immunofluorescence staining:

Anti-CK7 rabbit, EPR17078 (Abcam, 1:250)
Anti-GFP chicken, polyclonal (Invitrogen A10262, 1:250)
Anti-NGFR-APC, ME20.4–1.H4 (Miltenyi, 1:100)
Anti-Rabbit IgG Alexa Fluor (AF)647 donkey, polyclonal (Invitrogen, 1:1000)

Anti-Chicken IgY AF488 goat, polyclonal (Invitrogen A11039, 1:1000)

Images were acquired using confocal microscope Leica TCS SP8 SMD FLIM Laser Scanning Confocal at 20× magnification. Images were analyzed using ImageJ.

Images stained with anti-GFP primary antibody, that were subsequently stained with anti-HNF4α or anti-HNF1β antibodies for immunohistochemistry, were imaged on an ImageXpress Confocal MC system (Molecular Devices, San Jose, CA, USA) with a PlanApoLambda 20X/0.75 for 2 channels: hoechst (nuclei), 488 (GFP). A grid with multiple images was collected for individual tissues in order to cover the entire section. These images were then tiled using the software ArivisVision4D (ZeissAG, Germany) and co aligned with immunohistochemical images. Each individual bile duct comprising a GFP-positive BEC was then identified manually on each separate image and compared. In each bile duct GFP-positive BEC were counted as positive (GFP-positive and HNF4α- or HNF1β-positive, respectively) or negative (GFP-positive and HNF4α- or HNF1β-negative, respectively).

Immunohistochemistry imaging

Livers were perfused from the inferior vena cava with PBS EDTA 0.5mM to remove blood, then harvested and fixed at least 24h in zinc-formalin. Embedding and cutting of livers and slides preparation, staining and acquisition were performed by Centro di Imaging Sperimentale facility in San Raffaele hospital. Briefly, formalin-fixed paraffin-embedded consecutive sections (4 μm) were dewaxed and hydrate through graded decrease alcohol series and stained for histology or immunohistochemical characterization (IHC). Slides were immunostained with Automatic Leica BOND RX system (Leica Microsystems GmbH, Wetzlar, Germany). First, tissues were deparaffinized and pre-treated with the Epitope Retrieval Solution (ER1 Citrate Buffer for GFP staining or EDTA buffer for HNF4α or HNF1β staining). Anti-GFP (Invitrogen, A11122, 1:1,000), anti-mouse HNF4α (Cell Signaling Technologies, C11F12, 1:100) or anti-mouse HNF1β (Proteintech, polyclonal, 1:1,000) primary antibodies were incubated 30 min at room temperature and was developed with Bond Polymer Refine Detection (Leica, DS9800). Slides were acquired with Aperio AT2 digital scanner at magnification of 20X (Leica Biosystems) and analyze with Imagescope (Leica Biosystem).

Histology of organoids

Organoids were centrifuged at 1100 rpm for 5 min, then washed with PBS, fixed in buffered formalin for 1 h and washed with PBS. Fixed organoids were then plunged into HISTOGEL (Epredia) to obtain a cell-block which was further embedded in paraffin. 3-μm paraffin sections were cut and stained with haematoxylin and eosin (Bio-optica) for histological analysis.

ELISA

The concentration of human FVIII was determined in organoid supernatant by an enzyme-linked immunosorbent assay (ELISA) specific for human FVIII antigen. Microtiter plates were coated with anti-hFVIII binding Ab (Green Mountain Antibodies #GMA8016, 0.2 μg/well in 0.1 M carbonate buffer, pH 9.6) over night at 4°C and then blocked 1 h at room temperature with blocking buffer (PBS 0.05% Tween 20, 1M NaCl, 10% heat inactivated horse serum, Gibco). Samples are diluted as needed starting from 1:10 in blocking buffer, added to wells (100μL/well) and incubated 2 h at 37°C. hFVIII was detected by adding detection Ab (Affinity Biologicals, F8C-EIC-D) 1 h at 37°C, followed by 5–10 min incubation with 100μL/well of TMB substrate (Surmodics). Reaction was blocked with HCl 1N (50μL/well) and absorbance of each sample was determined spectrophotometrically at 450nm, using a Multiskan GO microplate reader (Thermo Fisher Scientific) and normalized to antigen standard curve (ReFACTO, Pfizer, from 25 ng/mL to 0.39 ng/mL serially diluted 1:2 in blocking buffer). The concentration of human FIX was determined in organoid supernatant or in the plasma of organoid transplanted mice by an ELISA specific for human FIX antigen. Microtiter plates were coated with anti-hFIX binding Ab (Haematologic Technologies #AHIX-5041, 0.45 μg/well in 0.1 M carbonate buffer, pH 9.6) over night at 4°C and then blocked 1 h at room temperature with blocking buffer (PBS 0.05% Tween 20, 1.5% bovine serum albumin). Samples are diluted as needed starting from 1:10 in blocking buffer, added to wells (100μL/well) and incubated over night at 4°C. hFIX was detected by adding detection Ab (1:20,000, Cederlane, CL20040APHP) 1 h at 37°C, followed by 5–10 min incubation with 100μL/well of TMB substrate (Surmodics). Reaction was blocked with HCl 1N (50μL/well) and absorbance of each sample was determined spectrophotometrically at 450nm, using a Multiskan GO microplate reader (Thermo Fisher Scientific) and normalized to antigen standard curve (BeneFIX, Pfizer, from 12.5 ng/mL to 0.19 ng/mL serially diluted 1:2 in blocking buffer supplemented with 20% murine plasma when murine samples were tested). The concentration of murine albumin was determined in organoid supernatant by an ELISA specific for murine albumin antigen (Mouse albumin ELISA Kit, Abcam) following manufacturer's instructions. Absorbance of each sample was determined spectrophotometrically, using a Multiskan GO microplate reader (Thermo Fisher Scientific) and normalized to antigen standard curves. The concentration of NHP albumin was determined in organoid supernatant or in mouse plasma by an ELISA specific for monkey albumin antigen (AssayMax monkey albumin ELISA Kit, AssayPro) following manufacturer's instructions. Absorbance of each sample was determined spectrophotometrically, using a Multiskan GO microplate reader (Thermo Fisher Scientific) and normalized to antigen standard curves.

Statistical analysis

Statistical analyses were performed upon consulting with professional statisticians at the San Raffaele University Center for Statistics in the Biomedical Sciences (CUSSB). When normality assumptions were not met, non-parametric statistical tests were

performed. Two-tailed Mann-Whitney test was performed to compare two independent groups, while in presence of more than two independent groups Kruskal-Wallis test followed by post-hoc analysis (Dunn's test for multiple comparisons against the reference control group along with Bonferroni's correction) was applied. For repeated measures over time, two-way ANOVA was performed. For paired observations, Wilcoxon matched-pairs signed rank test or Friedman test with Dunn's multiple comparisons test were performed. Inferential techniques were applied in presence of adequate sample sizes ($n \geq 5$), otherwise only descriptive statistics are reported.